Identification of Native Protein of a Novel Peroxisome Proliferator-Activated Receptor Alpha (PPARα) Target Gene – PPARα-Regulated and Starvation Inducible Gene (PPSIG) by Production of Polyclonal Antisera

YAU Wing Yiu, Winifred
B. Sc (Hons), CUHK

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Philosophy in Biochemistry

October 2006

Thesis/Assessment Committee

Professor C. C. Wan (Chair)

Professor S. S. T. Lee (Thesis Supervisor)

Professor C. C. Wan (Committee Member)

Professor K. B. Wong (Committee Member)

Professor W. K. Wong (External Examiner)

I declare that the assignment here submitted is original except for source material explicitly acknowledged. I also acknowledge that I am aware of University policy and regulations on honesty in academic work, and of the disciplinary guidelines and procedures applicable to breaches of such policy and regulations, as contained in the website http://www.cuhk.edu.hk/policy/academichonesty/

Signature | 31-10-2006 Date

YAU WING YIU WINIFRED | 04441250

Name | Student ID

BCH 604T | RESEARCH

Course code | Course title

Abstract

Peroxisome proliferator-activated receptors (PPARs) are nuclear transcription factors eliciting a variety of cellular responses by regulating a combination of target genes. Among the three isoforms, PPARα, PPARβ/σ and PPARγ, PPARα is well-characterized to play important roles in lipid metabolism, as well as induction of liver cancer when activated by peroxisome proliferators such as Wy-14,643 ([4-chloro-6(2,3-xylidino)-2-pyrimidinylthio]acetic acid), clofibrate and plasticizers.

A novel PPARα target gene - PPARα-regulated and Starvation Inducible Gene (PPSIG) was identified by fluorescent differential display in our laboratory. Its functions remain elusive while it shows sequence homology to phytoene dehydrogenase and is recently found to be a saturase. PPSIG is constitutively expressed in liver and its expression is dramatically induced by both 72-h starvation and 2-wk treatment with Wy-14,643 in PPARα wild-type mice but not in PPARα knockout mice. Open reading frame analysis suggests that PPSIG is a ~67 kDa protein. According to computer-assisted prediction and confocal microscopic observation of mammalian cells transiently over-expressing GFP-PPSIG fusion proteins, PPSIG is localized in the endoplasmic reticulum.

As PPSIG is a novel PPARα target gene, characterization of its function and biological role will be of high scientific significance. To start with, we investigated its subcellular localization and protein induction level in mouse liver. In order to identify the PPSIG protein in Western blottings, antibody specific to PPSIG was produced. A truncated length of PPSIG cDNA was first cloned into two prokaryotic expression vectors, pThioHis and pTYB and transformed into *E. coli* for the production of two recombinant proteins having different fusion partners. In order to minimize bacterial contaminants, the fusion proteins were semi-purified by cutting out the protein band from a preparative SDS-PAGE gel. Rabbits were immunized with 5 boostings of homogenized protein bands containing the fusion protein antigens. Using the polyclonal antisera obtained as primary antibodies in Western blottings of subcellular fractions, PPSIG was detected as a ~67 kDa protein localized primarily in the endoplasmic reticulum. The native PPSIG was immunoprecipitated from liver microsomes by the antiserum and identified by mass spectrometry. PPSIG protein induction analysis using liver microsomes showed that it was highly induced by 72-h starvation but only slightly induced by 2-wk Wy-14,643 treatment. The differentiated protein induction pattern suggested that starvation is more likely the physiological inducer of PPSIG protein, with the implication that PPSIG plays more a biological role than a toxicological role in mouse liver. The findings inevitably

facilitate experimental designs and help direct future researches on the functional studies of PPSIG.

摘 要

過氧化物酶體增殖因子激活受體 (PPARs) 屬於核轉錄因子，通過調控其靶基因而參與很多細胞應答途徑。在 PPARs 的三種亞型，PPARα, PPARβ/σ 和 PPARγ 中，對 PPARα 的研究是最透徹的。 PPARα 在脂質代謝中有重要的作用；同時在受到過氧化物酶體增殖因子例如 Wy-14,643，安妥明和可塑劑的激活時， PPARα 可以誘發肝臟腫瘤。

一個新的 PPARα 靶基因 － PPARα 調控且饑餓誘導基因 (PPSIG) 是本實驗室通過熒光差異顯示的方法篩選到的。 PPSIG 的功能還沒有明確，特別是因爲它的序列與八氫番茄紅素脫氫酶同源而又有證據顯示它是一個飽和酶。 PPSIG 在肝臟組織中表達；在饑餓七十二小時或喂食 Wy-14,643 兩周後， PPSIG 在野生型小鼠肝臟中的表達顯著提高，而在 PPARα 缺失的小鼠中沒有變化。開放閱讀框架分析顯示 PPSIG 是一個約 67 kDa 的蛋白。根據電腦輔助預測和利用共聚焦顯微鏡觀測瞬時轉染並高表達 GFP-PPSIG 融合蛋白的哺乳動物細胞， PPSIG 被定位於細胞內質網。

由于 PPSIG 是一個新的 PPARα 靶基因，對于其功能以及生物學作用的研究是很有科學意義的。本論文中，我們將著重研究 PPSIG 的亞細胞定位以及其蛋白質在小鼠肝臟中表達水平的誘導變化。爲了在 Western 雜交中檢測 PPSIG 蛋白，我們製備了特異性的 PPSIG 抗體。PPSIG 的 cDNA 片段被克隆

到 pThioHis 和 pTYB 兩個不同的原核表達載體中，然後被轉化入大腸杆菌并表達兩個不同的融合蛋白。爲了減少細菌污染，通過對製備型 SDS-PAGE 膠進行切割而分離 PPSIG 融合蛋白條帶。對白兔進行五次增强免疫，每次注射經過勻漿的含有 PPSIG 融合蛋白抗原的膠條。在 Western 雜交中用 PPSIG 多克隆抗血清作爲第一抗體檢測亞細胞組分，初步檢測到細胞內質網中有約 67 kDa 的 PPSIG 蛋白表達。利用抗血清對小鼠肝臟微粒體進行免疫沉澱，分離到了內源性 PPSIG 并通過了質譜儀的鑒定。對肝臟微粒體進行 PPSIG 誘導表達分析，饑餓七十二小時可以誘導 PPSIG 的高水平表達而喂食兩周的 Wy-14,643 對 PPSIG 的誘導只是很輕微的。這樣不同的 PPSIG 誘導表達圖譜提示饑餓條件更可能是 PPSIG 蛋白的生理學誘導因子，從而說明 PPSIG 在小鼠肝臟中更主要的是生物學功能而不是其毒物學功能。這一發現有利於推動實驗決策並且幫助確立進一步的對 PPSIG 功能研究的方向。

Acknowledgement

I would like to express my heartfelt gratitude towards my thesis supervisor, Professor Lee Sau Tuen Susanna for her insightful guidance and advice on my project, as well as her trust and understanding during times of obstacles. Special thanks go to Professor Cheung Wing Tai for his enthusiastic advice and suggestions. The two years of M. Phil studies would be a valuable enlightenment in my future career of Biochemistry research.

I would like to thank all my labmates in SC 294, Carol Chan, Yo Yuen, Sun Yan, Lui Ng, Alice Leung, May Li, Kenny Ng, Wolf, Tom, Ho, Jeremy Yeung and Bonita Chan for their company during times of happiness and difficulties. I am also grateful to have the presence of my friends, Tsang Tsun Yee, Carmen Hon, Lydia Lung and Gladys Teng to share my feelings.

Last but not least, I would like to thank my beloved parents and brother who support me always with endless love in God.

Table of Contents

List of Abbreviations

BCIP	5-bromo-4-chloro-3-indolyl-phosphate
BLAST	Basic Local Alignment Search Tool
bp	Basepair
cm	Centimeter
cDNA	Complementary deoxyribonucleic acid
CTL	Control diet
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FDD	Fluorescent differential display
g	Gram
h	Hour
HCl	Hydrochloric acid
IPTG	Isopropyl-beta-D-thiogalactopyranoside
LB	Luria broth
Kb	Kilobase
KO	Knockout
mA	Milliampere
μg	Microgram
mg	Milligram
min	Minute
μl	Microlitre
ml	Millilitre
μM	Micromolar
mM	Millimolar
NaCl	Sodium chloride
NBT	Nitroblue tetrazolium chloride
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
PP	Peroxisome proliferator
PPAR	Peroxisome proliferator-activated receptor
PPARα	Peroxisome proliferator-activated receptor alpha
PPARβ	Peroxisome proliferator-activated receptor beta

PPARγ	Peroxisome proliferator-activated receptor gamma
PPRE	Peroxisome proliferator responsive element
PPSIG	Peroxisome proliferator and starvation inducible gene
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyarcylamide gel electrophoresis
Tris-HCl	Tris (hydroxymethyl)aminomethane/hydrochloric acid
V	Voltage
Wy	Wy-14,643 [4-chloro-6(2,3-xylidino)-2-pyrimidinylthio]acetic acid
Wy-14,643	[4-chloro-6(2,3-xylidino)-2-pyrimidinylthio]acetic acid

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Peroxisome proliferator-activated receptors (PPARs)

1.1.1 What are PPARs?

PPARs belong to the nuclear hormone receptor superfamily, which also includes receptors for retinoic acid, thyroid hormone and estrogen. They are single polypeptide proteins with a variable amino-terminal domain, a highly conserved DNA-binding domain and a carboxyl-terminal ligand-binding domain (Ribeiro *et al.* 1995). They are regulated by exogenous and endogenous ligands known as peroxisome proliferators (PPs), which bind to the receptor and activates its translocation to the nucleus. After formation of a heterodimer with 9-*cis*-retinoid X receptor (RXR), PPARs bind to their corresponding responsive elements and regulate the expression of their target genes by either inducing or inhibiting their transcription (Reddy and Hashimoto 2001).

1.1.2 PPAR ligands – peroxisome proliferators

Peroxisome proliferators are a diverse group of chemicals that causes peroxisome proliferation in rodent liver, characterized by liver enlargement due to

hyperplasia and hypertrophy (Dzhekova-Stojkova *et al.* 2001). A group of potent peroxisome proliferators was hypolipidemic drugs, e.g. Wy-14,643, clofibrate and ciprofibrate (Lock *et al.* 1989). Besides hypolipidemic drugs, plasticizers such as diethylhexyl phthalate also causes certain extent of peroxisome proliferation. Other PPs include environmental contaminants, e.g. herbicides, industrial solvents and food flavours (Dzhekova-Stojkova *et al.* 2001). Structure of PPs usually consists of a carboxylic functional group and a large hydrophobic domain, which is similar to fatty acids. In fact, fatty acids are endogenous PPs that can be induced under various physiological conditions such as starvation, high fat diets and diabetes mellitus (Vanden Heuvel 1999). Starvation usually causes differential expression of peroxisomal enzymes in similar manner to exogenous PPs (Leone *et al.* 1999).

1.1.3 PPAR isoforms

PPARs exist in three isoforms, PPARα, PPARβ/σ and PPARγ. Encoded by different genes, they exhibit different tissue distribution and function. PPARα is predominantly expressed in liver and intestine and participates actively in energy metabolism (Ehrmann *et al.* 2002; Reddy and Hashimoto 2001). PPARβ/σ is expressed at a slightly higher level in intestine than in liver and spleen (Ehrmann *et al.* 2002). Its physiological role is relatively less characterized, but it has been found to

induce differentiation of epithelial cells while inhibiting cell proliferation (Burdick *et al.* 2006; Michalik *et al.* 2003). PPARγ is localized mostly in adipocytes and is also found in spleen and intestine (Ehrmann *et al.* 2002). It plays an important role in adipocyte differentiation and takes part in insulin sensitization, anti-inflammatory and anti-atherogenic effects (Haag and Dippenaar 2005; Lehrke and Lazar 2005; Madsen *et al.* 2005; Racke *et al.* 2006; Pineda Torra *et al.* 1999). Although the three PPAR isoform differ in ligand specificity and tissue distribution, their physiological roles are interrelated, accounting for their pleiotropic effects.

1.2 Biological roles of PPARα

1.2.1 Lipid metabolism

Extensive studies have focused on the role of PPARα on lipid metabolism. A number of PPARα target genes have been identified to play important roles in fatty acid uptake and β-oxidation in mitochondria and peroxisomes (Mandard *et al.* 2004). Acyl-CoA dehydrogenase in mitochondria oxidizes short to long chain fatty acids (Djouadi *et al.* 2005) while peroxisomes contains acyl-CoA oxidase that oxidizes very long chain fatty acids (Chu *et al.* 1995). During starvation, PPARα masters the change in energy source by regulating the expression of cytochrome P450 4A enzymes involved in microsomal fatty acids ω-hydroxylation (Kroetz *et al.* 1998) and

mitochondrial 3-hydroxy-3-methylglutaryl-CoA (mHMG-CoAS), a rate-limiting enzyme in ketogenesis (Rodriguez *et al.* 1994). As a counter effect to fatty acid oxidation, PPARα was also reported to induce lipogenesis by upregulate a number of lipogenic enzymes, e.g. stearoyl-CoA desaturase-1 (SCD-1) and long chain free fatty acid elongase, a possible feedback mechanism to prevent excessive fatty acids breakdown (Zhou *et al.* 2006). In addition, it promotes plasma HDL and triglyceride clearance by upregulating the expression of apolipoproteins (Edvardsson *et al.* 2006) and help excretion of fatty acid by mediating hepatic bile acid metabolism (Barbier *et al.* 2003)

1.2.2 Glucose metabolism

PPARα plays a more important role in glucose metabolism during deprivation of glucose. It is thought to prevent decreased plasma glucose level by activating pathways for glucose production, as hypoglycemia is observed in PPARα knockout mice (Chakravarthy *et al.* 2005). During early stage of starvation, PPARα is believed to cause hepatic glycogen breakdown although no target gene has been identified. It also down-regulates the activity of phosphoenol-pyruvate carboxykinase that is responsible for gluconeogenesis (Srivastava *et al.* 2006). In addition, it reduces the utilization of glucose through glycolysis by activating

pyruvate dehydrogenase kinase isoform 4 (Sugden *et al.* 2001).

1.2.3 Inflammation

Although not as potent as PPARγ, PPARα exerts a certain effect on inhibiting inflammation, usually through trans-repression of related genes (Mandard *et al.* 2004). PPARα activation down-regulates the expression of endothelin-I, vascular cell adhesion molecule 1 (VCAM-1), IL-6 and cyclooxygenase-2 (Dyroy *et al.* 2005; Mandard *et al.* 2004; Pineda Torra *et al.* 1999; Racke *et al.* 2006; Zambon *et al.* 2006). By inhibiting the production of cytokines and growth factors, it inactivates endothelial cells, smooth muscle cells and macrophages, which are responsible for inflammatory response.

1.2.4 Oxidative stress

During fatty acid oxidation, the reaction catalyzed by the PPARα target gene acyl-CoA oxidase generates reactive intracellular peroxide (H_2O_2) molecules that can lead to cell damage. Accumulation of peroxide in cells results in oxidative stress, but it is inconclusive whether the level of oxidative stress produced by PPARα is enough to cause oxidative damage to macromolecules (Bosgra *et al.* 2005; Peters *et al.* 2005).

1.2.5 Cell proliferation and apoptosis

The mechanism of PPARα-mediated cell proliferation and apoptosis are not fully known, but may be contributed by the generation of intracellular free radicals (Bosgra *et al.* 2005; Peters *et al.* 2005). Studies have shown that the induction of DNA synthesis and suppression of apoptosis by peroxisome proliferators required tumour necrosis factor alpha receptor 1 (TNFα-1) and interleukin 1 (IL-1) receptor through the activation of MAK kinase signaling pathway. (Cosulich *et al.* 2000; Hasmall *et al.* 2000; West *et al.* 1999).

1.3 PPARα in health and diseases

1.3.1 Wound-healing

Activation of PPARα reduces apoptosis and lessens the extent of tissue injury. It inhibits proliferation of endothelial cells and macrophages, which causes inflammation. Activation of PPARα by its ligand lengthens the time to produce an inflammatory response, thereby giving more time for tissue recovery (Dyroy *et al.* 2005). PPARα is transiently induced after injury and it ligands have shown protective effect in injured kidney, intestine and lung (Michalik and Wahli 2006). The extent of tissue injury after ischemia/reperfusion was greater in PPARα knockout mice than wild-type mice.

1.3.2 Anti-atherogenesis

Early stage of atherosclerosis involves expression of VCAM-1 that plays a role in monocytes recruitment to atherosclerosis lesions. Monocytes, when activated by monocyte chemoattractant protein-1 (MCP-1) and interleukin 8 (IL-8), will differentiate into macrophages and form foam cells that eventually lead to plaques. By inhibiting cytokine expression and macrophage proliferation, PPARα helps prevent atherogenesis (Lefebvre *et al.* 2006; Pineda Torra *et al.* 1999).

1.3.3 Neuroprotection

PPARα was shown to preventively protect the brain from cerebral ischemia. Activation of PPARα improves sensitivity to endothelial relaxation, regulates oxidative stress by upregulating antioxidant enzymes, and decreases VCAM-1 and intercellular adhesion molecule 1 (ICAM-1) expression. The PPARα-dependence of the neuroprotective effect is yet to be confirmed and independent of PPARα's role in lipid metabolism (Deplanque *et al.* 2003).

1.3.4 Carcinogenesis

Long term treatment of Wy-14,643, a potent PPARα ligand, produced 100% incidence of liver cancer in PPARα wild-type but not knockout mice. Oxidative

stress is being considered a major cause for carcinogenesis because of its devastating ability. Free radicals produced can easily get into nucleus and induce DNA mutation (Bosgra *et al.* 2005). These free radicals activate the oxidative signaling pathway in parallel, which stimulates cell proliferation and inhibits apoptosis. It inhibits the programmed death of cells containing DNA with undesirable mutation. If these cells are further stimulated to replicate, they will soon be transformed into carcinoma (Peters *et al.* 2005).

The diverse effect of PPARα suggests that it plays a pivotal role in many important pathways, with the orchestration of different target genes. The identification of the whole spectrum of target genes will inevitably help unveiling the physiological and pharmacological role of PPARα.

1.4 PPARα-regulated and starvation inducible gene (PPSIG)

1.4.1 PPSIG is a PPARα target gene

A ~580 bp cDNA fragment (D3#3) was found to display at higher level in PPARα wild-type mice starved for 72 h in a fluorescent differential display, performed with AP1 and ARP3 primers (Lo, paper submitted). It was recovered from the gel, reamplified and used as a probe for Northern blotting using liver RNA

from PPARα wild-type and knockout mice either fed or starved for 72 h. The probe detected two transcripts at ~2000 and ~3000 bp, both showing a significantly higher expression level in wild-type mice starved for 72 h than the other three groups of mice (Ng, 2005). The same transcripts were detected with a higher expression level in wild-type but not knockout mice treated with Wy-14,643 for 11 months. As the gene can be induced by both PP and starvation, the gene was named PPARα-regulated and starvation inducible gene (PPSIG). The induction was only observed in PPARα wild-type but not knockout mice, further suggesting that PPSIG is a putative PPARα target gene. Recent discovery of a peroxisome proliferator responsive element (PPRE) in intron 1 (+375 to + 387) of PPSIG further supported its regulatory effect by PPARα (Ng, 2005).

1.4.2 Computer-assisted predictions on PPSIG

Full length of PPSIG mRNA was obtained from the FDD fragment by rapid amplification of cDNA ends (RACE). By submitting the mRNA sequence of PPSIG for BLAST searching, it was found to share 100% identity with RIKEN cDNA 0610039N19 gene with accession number 87044905. As the RIKEN gene was a novel gene with unknown function, the translated PPSIG protein sequence was subjected to putative conserved domain analysis, and the protein was predicted to

belong to a family of phytoene dehydrogenase having a domain with amino oxidase activity. Protein sequence analysis using TMHMM2.0 software showed that PPSIG may have a transmembrane segment containing amino acid residues 5-27. Subcellular localization prediction by WoLF PSORT software also showed that PPSIG is a membrane-bound protein localized in the endoplasmic reticulum. For signal peptide analysis using SignalP 3.0 Server, PPSIG is predicted to contain a secretary signal peptide with residues 1-25, which may be cleaved off upon secretion.

1.4.3 Current characterization of PPSIG

In 2004, a group of researchers working on retinoids identified the RIKEN gene as all-*trans*-13,14-dihyroretinol saturase (RetSat) (Moise *et al.* 2004). RetSat was reported to catalyze the *in vitro* conversion of all-*trans*-retinol to all-*trans*-13,14-dihydroretinol with no PPARα dependence described and biological role yet unknown. PPSIG mRNA has a high constitutive expression level in liver, kidney, small intestine, white fat and brown fat and can be induced in different organs, markedly in liver and kidney (Lo, 2003). It was found to localize in the endoplasmic reticulum by observing the compartmentation of GFP-PPSIG fusion protein in mammalian cells (Sun, paper submitted).

1.5 Objectives of the present study

PPSIG is at present a novel PPARα target gene. Further characterization such as proteomics and functional studies will be needed to unveil its biological role. In order to identify of PPSIG protein from tissue samples, the current study aims to produce a polyclonal antiserum immunogenic to PPSIG.

The project includes several objectives:

(1) To clone a partial length of PPSIG into pThioHis and pTYB expression vectors for the expression of Thio-PPSIG and Intein-PPSIG fusion proteins,

(2) To immunize rabbits with the PPSIG fusion protein antigens for generation of polyclonal antisera,

(3) To test for antigenicity of the antisera against its bacterial recombinant antigen on Western blottings,

(4) To identify the native PPSIG from liver microsomes assisted by protein induction by starvation and Wy-14,643 in PPARα wild-type and knockout mice,

(5) To confirm the native PPSIG by mass spectrometry followed by immunoprecipitation.

Chapter 2 Materials and Methods

2.1 Materials

Wy-14,643 was purchased from Chemsyn Science Laboratories (Lenexa, KS, USA). Mouse chows containing 0.0% or 0.1% (w/w) Wy-14,643 were produced by Bioserv Company (Frenchtown, NJ, USA). 100 bp and 1 kb DNA ladders, Benchmark™ prestained protein ladder, polyclonal anti-PMP70 antibody and monoclonal anti-thio antibody were obtained from Invitrogen (Carlsbad, CA, USA). Nuclease-free water (NF water) and supercoiled DNA ladder was obtained from Promega (Madison, WI, USA). DIG-labeled RNA molecular weight marker I was purchased from Roche (Germany). SDS-PAGE natural standards, Kaleidoscope prestained standards and prestained standards were obtained from Bio-Rad Laboratories (Hercules, CA, USA). Restriction enzymes and polyclonal anti-(chitin binding domain) antiserum (anti-CBD) were purchased from New England Biolabs (Beverly, MA, USA). Bacto™ tryptone and yeast extract were purchased from BD Biosciences (San Jose, CA, USA). Polyclonal anti-calnexin antibody and anti-cpn10 antibody were purchased from Stressgen (Ann Arbor, MI, USA). All other chemicals of the highest analytical grade were purchased from Sigma (St. Louis, MO, USA) and USB chemical (Cleveland, OH, USA).

2.2 Animals and treatment

PPARα knockout mice generated on SV/129 background was obtained from the US National Cancer Institute (National Institutes of Health, Bethesda, MD, USA) (Lee *et al.* 1995). They were inbred and reared in the Laboratory Animal Services Centre of the Chinese University of Hong Kong together with PPARα wild-type mice (SV/129 mice) on a 12-h light/dark cycle. For starvation experiment, three-month old male PPARα wild-type and knockout mice were starved for 3 days while control group mice were fed with normal diet (Laboratory Diet™ brand mouse diet No. 5015, PMI Nutrition, St. Louis, MO). For treatment with Wy-14,643, mice were fed with mouse chow containing either 0.0% (control group) or 0.1% (w/w) Wy-14,643 for 2 weeks. Mice were sacrificed by cervical dislocation. Their livers were collected and frozen immediately in liquid nitrogen. Half of each liver was used for total RNA extraction and half for preparation of microsomes.

2.3 Cloning of PPSIG into pThioHis and pTYB expression vectors

2.3.1 PCR amplification of PPSIG cDNA insert

2.3.1.1 PPSIG cDNA insert for pThioHis vector

PPSIG was cloned into a pThioHis C expression vector (Figure 2.1) using the

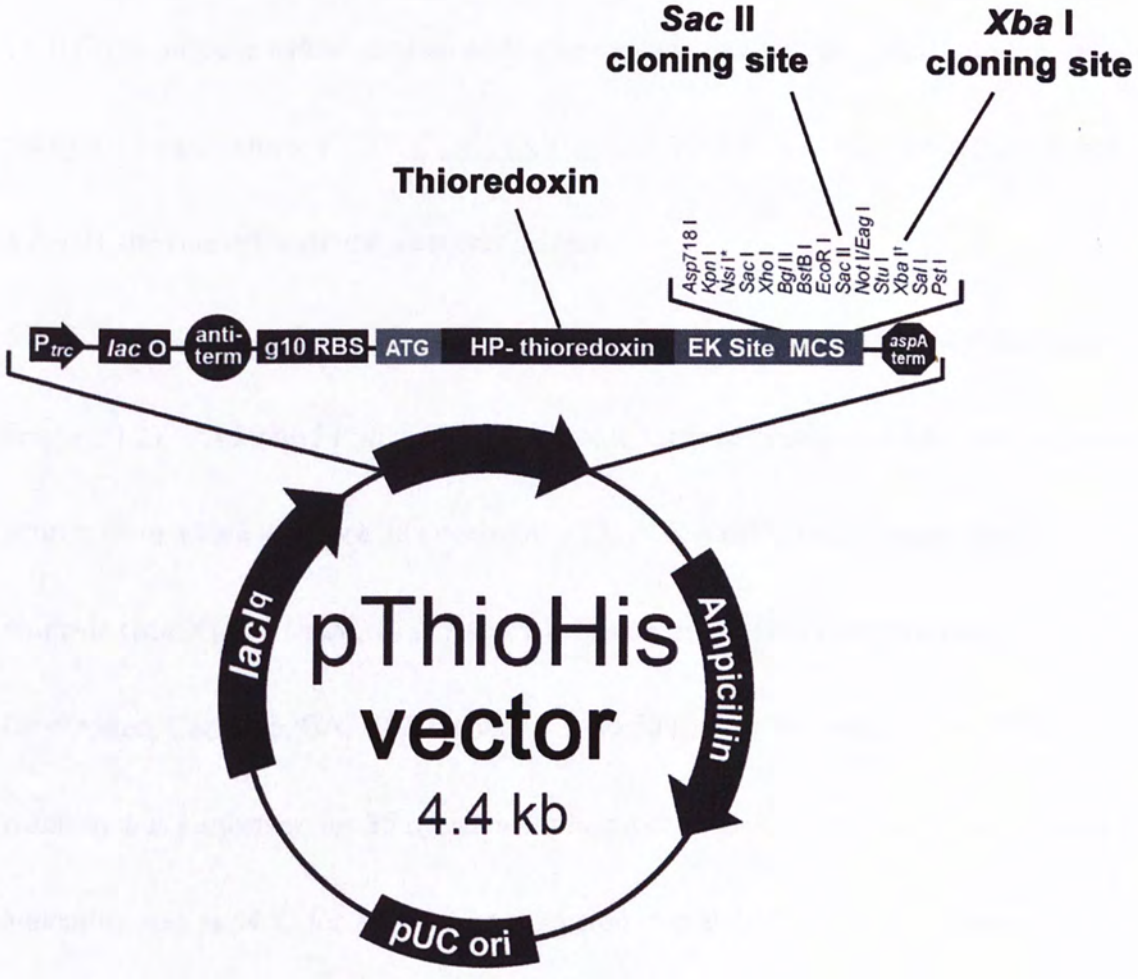


Figure 2.1 Map of pThioHis vector

The pThioHis vector allowed the addition of the thioredoxin fusion partner at the N-terminal of PPSIG. The *Sac* II and *Xba* I digestion sites were chosen for cloning of PPSIG cDNA.

Invitrogen His-Patch ThioFusion Expression System (Carlsbad, CA, USA). A partial length of PPSIG inclusive of the stop codon was amplified from a RACE clone (5'II32) to produce a PCR product of 811 bp containing *Xba* I and *Sac* II cloning sites using a forward primer 5'-TCCCCGCGGGGAATGCGGGAATGTTCAATAC-3' with a *Sac* II site (underlined) and a reverse primer 5'-GCTCTAGAGCTTCTCTGAACGGACTACAT-3' with an *Xba* I site (underlined) (Figure 2.2). A 20 ng PPSIG cDNA, 10 µmole forward primer and 10 µmole reverse primer were added to a core mix containing 1X PCR buffer, 2 mM magnesium chloride ($MgSO_4$), 0.1 mM dNTPs and 1 U Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA, USA), making up to 50 µl with NF water. The PCR reaction was performed for 35 cycles including a denaturing step at 94°C for 30 s, an annealing step at 54°C for 30 s and an extension step at 68°C for 2 min using a thermocycler (Applied Biosystem GeneAMP PCR 9700). DNA concentrations were measured by absorbance at 260 nm using a scanning spectrophotometer (Beckman DU640 UV-VIS-NIR).

2.3.1.2 PPSIG cDNA insert for pTYB vector

PPSIG was cloned into a second pTYB1 expression vector (Figure 2.3) using the New England Biolabs Impact™-CN protein fusion and purification system

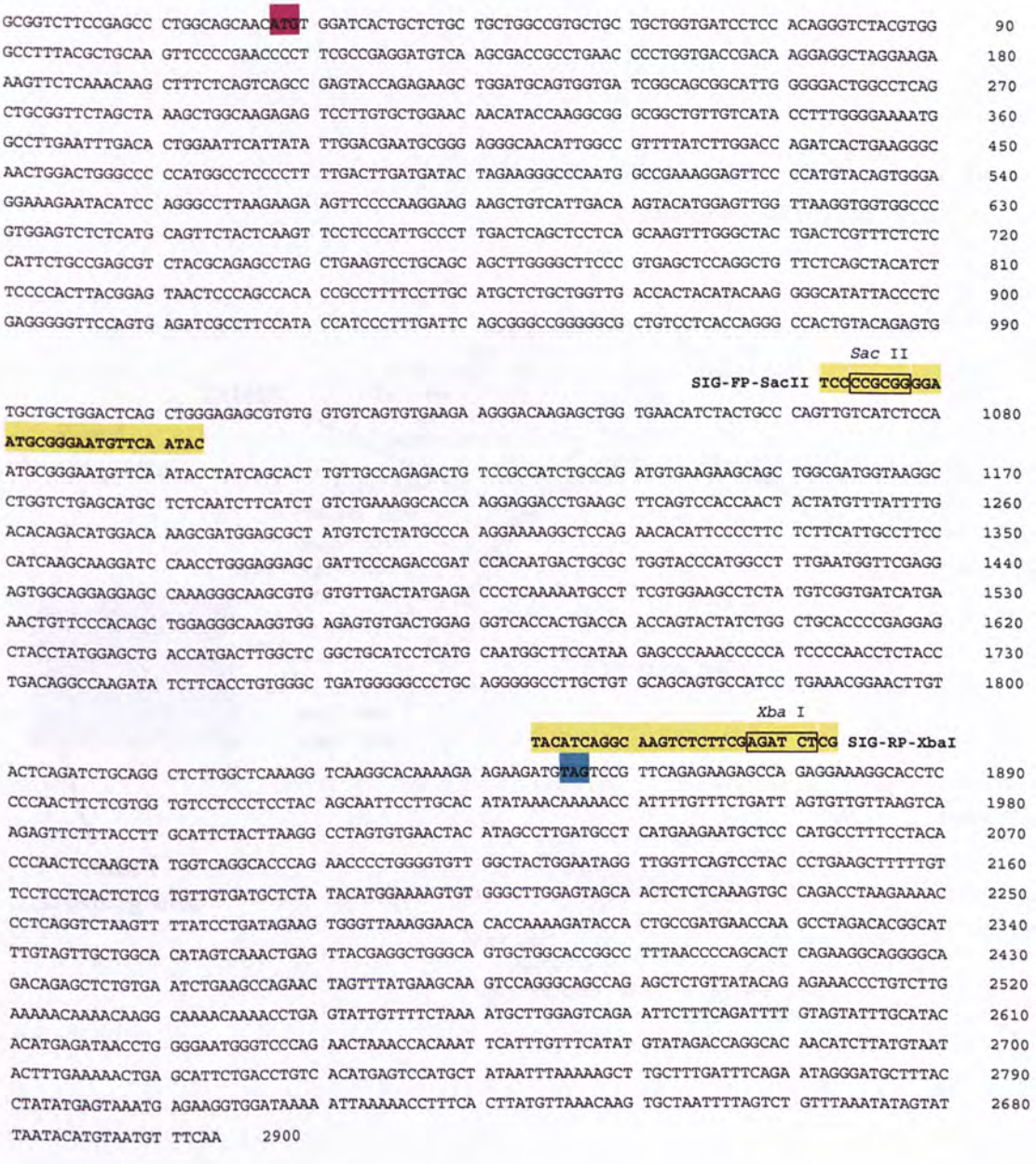


Figure 2.2 Primer design for PCR amplification of PPSIG containing *Xba* I and *Sac* II sites for cloning into pThioHis vector

A set of primers (highlighted in yellow) was designed to amplify a partial PPSIG cDNA of 811 bp, with the *Sac* II digestion site-containing forward primer located at 1081 bp of the full length PPSIG cDNA and the *Xba* I site-containing reverse primer located at 1856 bp region inclusive of the stop codon of PPSIG (highlighted in blue). The start codon of the full length PPSIG was highlighted in red.

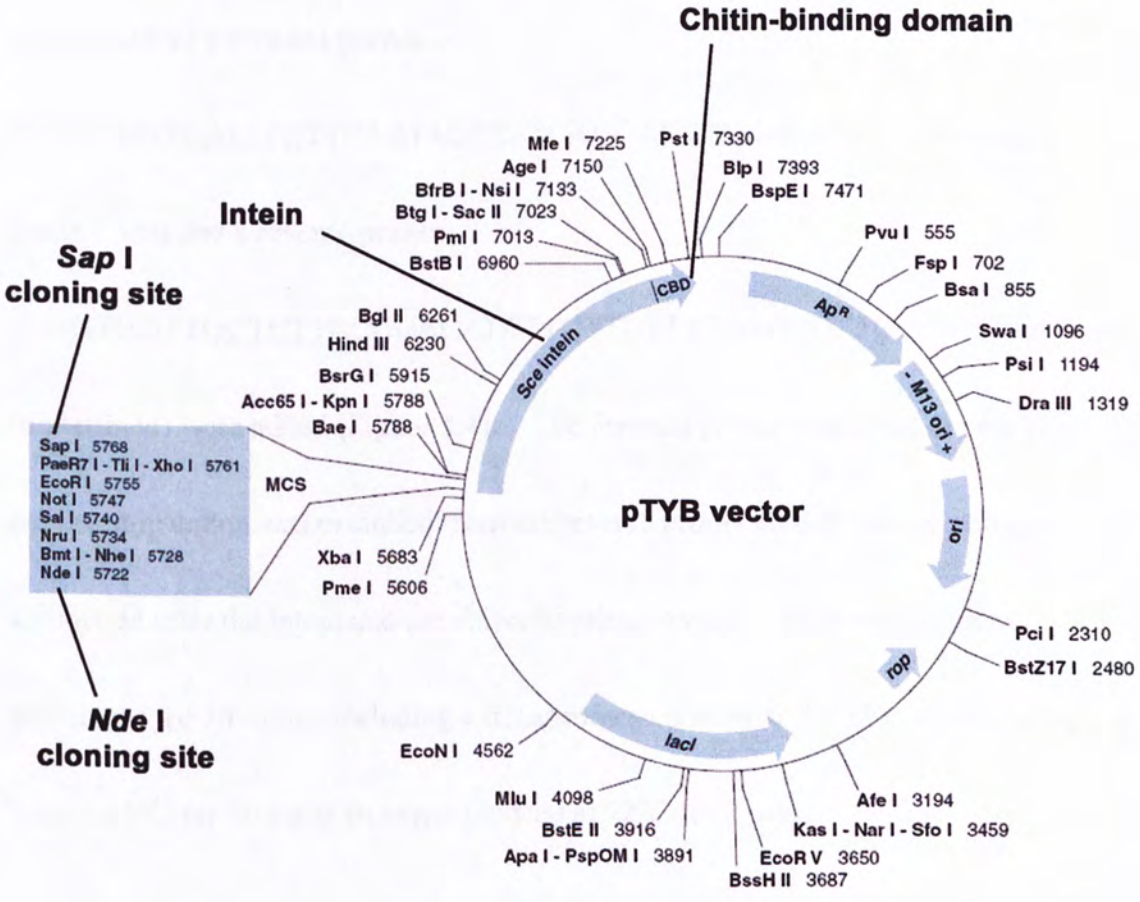


Figure 2.3 Map of pTYB vector

The pTYB vector allowed the addition of intein at the C-terminal of PPSIG. Chitin binding domain (CBD) is present at the C-terminal of intein to facilitate fusion protein detection by Western blotting. The *Sap* I and *Nde* I digestion sites were chosen for cloning of PPSIG cDNA.

(Beverly, MA, USA). The partial PPSIG was amplified by PCR to produce a cDNA of 792 bp using the same PCR mixture as described in section 2.3.1.1 except that 10 μmole each of a forward primer

5'-GGTGGTCATATGTTCAATACCTATCAGCACT-3' with *Nde* I digestion site (underlined) and a reverse primer

5'-GGTGGTTGCTCTTCCGCACATCTTCTTCTTTTGTGCCTTG-3' with *Sap* I site (underlined) were added (Figure 2.4). The forward primer contained the start codon, but the stop codon was excluded from the reverse primer so that transcription is attenuated after the intein and the chitin-binding domain. PCR reaction was performed for 30 cycles including a denaturing step at 94°C for 30 s, an annealing step at 65°C for 30 s and an extension step at 72°C for 2 min.

2.3.2 Restriction enzyme digestion of PPSIG cDNA insert and pThioHis vector

The PPSIG cDNA insert and the expression vector pThioHis C were double-digested with *Xba* I and *Sac* II. Digestion mixture of cDNA insert contained 2 μg cDNA, 20 U *Sac* II, 20 U *Xba* I, 1X NEBuffer 2 and 1X BSA added up to 40 μl with autoclaved distilled water. Digestion mixture of pThioHis vector contained 5 μg DNA, 60 U *Sac* II, 60 U *Xba* I, 1X NEBuffer 2, 1X BSA and autoclaved distilled

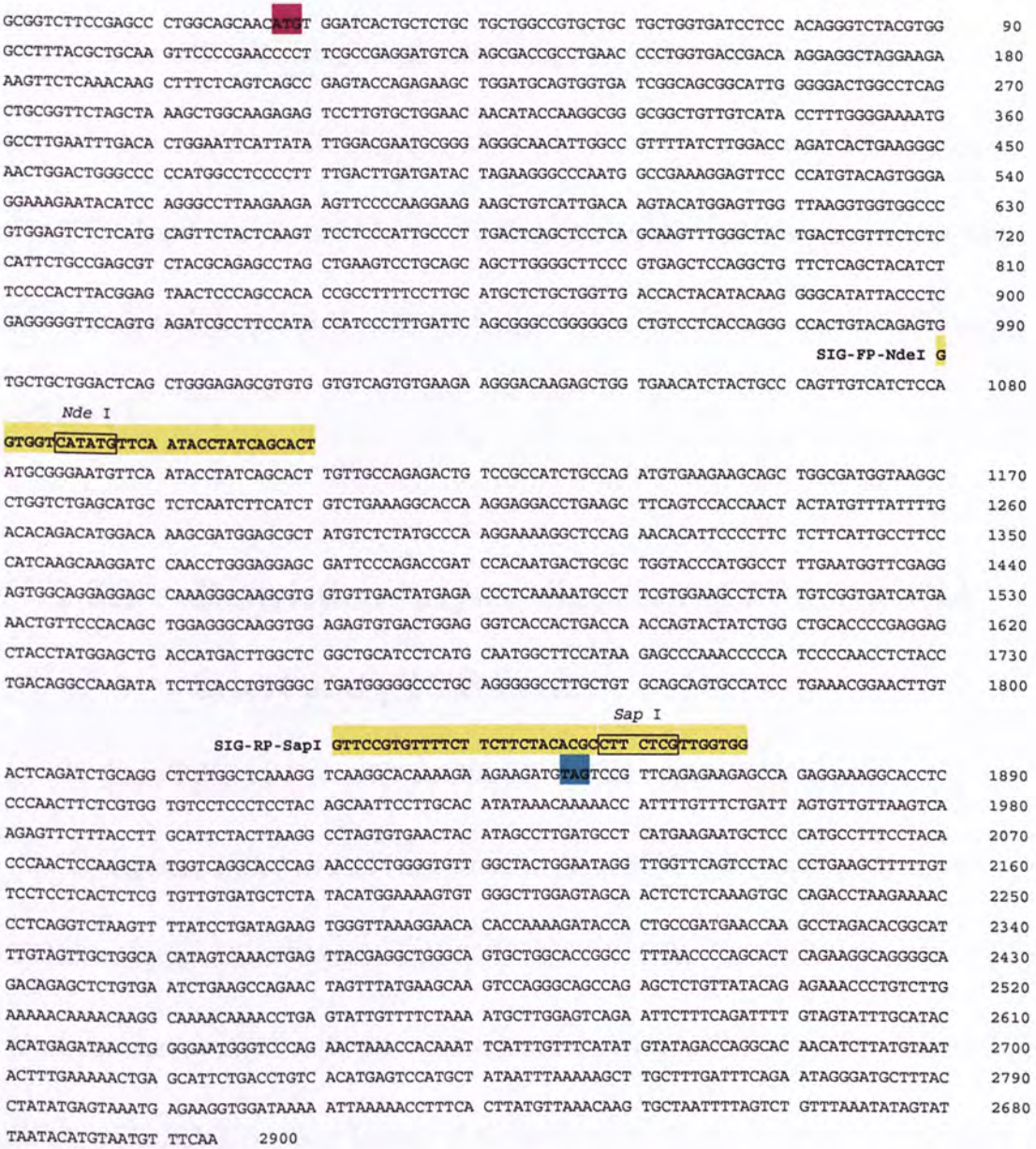


Figure 2.4 Primer design for PCR amplification of PPSIG containing *Nde* I and *Sap* I sites for cloning into pThioHis vector

A set of primers (highlighted in yellow) designed to amplify a partial PPSIG cDNA of 792 bp included a forward primer located at 1089 bp of the full length PPSIG cDNA containing the *Nde* I digestion site and a reverse primer located at 1853 bp region containing the *Sap* I site that excluded the stop codon (highlighted in blue). The start codon of the full length PPSIG was highlighted in red.

water with a total volume of 50 µl. Both digestion mixtures were kept in a 37°C water bath overnight. The digested DNA was resolved on 1% or 1.5% agarose gels. The DNA band was visualized on a UV Transilluminator (UVP M20), cut out with a razor blade and gel-purified using the Qiagen QIAquick gel extraction kit (Hilden, Germany).

2.3.3 Restriction enzyme digestion of PPSIG cDNA insert and pTYB vector

The PPSIG cDNA insert and the expression vector pTYB were double-digested with *Nde* I and *Sap* I overnight at 37°C. Digestion mixture of cDNA insert contained 2 µg cDNA, 20 U *Nde* I, 3 U *Sap* I, and 1X NEBuffer 4 in 100 µl of autoclaved distilled water. Digestion mixture of pTYB vector contained 5 µg DNA, 60 U *Nde* I, 6 U *Sap* I, and 1X NEBuffer 4 in 100 µl of autoclaved distilled water. The digested DNA were resolved on 0.6% or 1.5% agarose gels, cut out with a razor blade and gel-purified using QIAquick gel extraction kit. DNA concentration was measured by absorbance at 260 nm.

2.3.4 Ligation and transformation

Ligation was carried out with digested PCR product and expression vector in a

3:1 molar ratio at 16°C overnight. For cloning into pThioHis vector, the ligation mixture contained 100 ng vector DNA, 52.5 ng insert cDNA, 0.5 U T4 DNA ligase and 1X ligase buffer in 10 µl NF water. For cloning into pTYB vector, the ligation mixture contained 100 ng vector DNA, 32 ng insert cDNA, 0.5 U T4 DNA ligase and 1X ligase buffer in 10 µl NF water. The ligation product was incubated on ice with 200 µl competent cells (TOP10 for pThioHis vector and HB101 for pTYB vector) before applying a 90 s heat shock at 42°C. After standing on ice for a further 2 min, 800 µl LB medium was added to the mixture and the culture was shaken at 250 rpm at 37°C for 1 h using an orbital shaker (Lab-Line). Cells were pelleted by centrifugation at 2,500 rpm for 8 min and 800 µl supernatant was removed. After resuspending in the remaining 200 µl supernatant, the transformed cells were spread on an LB agar plate containing 50 µg/ml ampicillin and allowed to grow at 37°C overnight.

2.3.5 Screening for recombinants by phenol/chloroform method

Clones containing recombinant plasmids were screened by phenol/chloroform method to check whether the plasmid had the expected size. A 20 µl of overnight culture of the bacterial colonies were added to 20 µl of phenol/chloroform and

vortexed for 1 min. After centrifugation at 14,000 rpm for 1 min at 4°C, the supernatant was loaded on a 0.7% agarose gel. Clones having plasmids of desired size were grown in 5 ml LB containing 50 μg/ml ampicillin and plasmid DNA was purified using the Mini-M® plasmid DNA extraction system (Viogene, Sunnyvale, CA, USA).

2.3.6 Confirmation of recombinant plasmid by restriction enzyme digestion

2.3.6.1 Digestion of pThioHis-PPSIG plasmid with *Xba* I and *Sac* II

Restriction enzyme digestion was performed to check if the plasmid carried PPSIG insert. A 0.5 μg plasmid DNA was digested with 20 U *Xba* I and 20 U *Sac* II, in 1X NEBuffer 2 and 1X BSA in 20 μl of autoclaved distilled water. After overnight incubation in a 37°C water bath, the digestion mixture was resolved on a 1% agarose gel.

2.3.6.2 Digestion of pTYB-PPSIG plasmid with *Eco*R V

Restriction enzyme digestion was performed using 0.5 μg plasmid DNA in a digestion mixture containing 10 U *Eco*R V, 1X NEBuffer 3, and 1X BSA in 20 μl of

autoclaved distilled water at 37°C overnight. The digestion mixture was resolved on a 0.7% agarose gel.

2.3.7 Transformation into expressing *E. coli* strains

After confirming the recombinant plasmid by restriction enzyme digestion, the plasmid was transformed into a suitable expressing *E. coli* strain with the same transformation procedure described in section 2.3.4. The pThioHis-PPSIG plasmid was transformed into BL21 star (DE3) competent cells and the pTYB-PPSIG plasmid into ER2566 competent cells for Thio-PPSIG and Intein-PPSIG fusion protein expression respectively.

2.4 Over-expression of PPSIG proteins in *E. coli*

Overnight bacterial culture containing the desired DNA plasmid was inoculated in 100 volume of fresh LB medium (1% bacto-yeast extract, 0.5% bacto-tryptone and 1% NaCl) for 37°C incubation with orbital shaking in a Lab-line Environ Shaker until absorbance at 600 nm reached 0.6. Protein induction was carried out at 37°C for 4 h in the presence of 1 mM IPTG. Bacterial cells were pelleted at 6,000 g for 15 min and resuspended in 1/10 volume of wash buffer (2 mM Tris-HCl, pH 7.5, 10 mM EDTA and 1% Triton X-100). Lysozyme was added to the

resuspension to a final concentration of 0.2 mg/ml for a 15-min incubation at 30°C. Proteins were released by sonication for 3 x 10 s using the Ultrasonic Processor (DAIGGER GE130PB). Insoluble proteins were expressed as inclusion bodies and were collected by centrifugation at 10,000 g for 10 min. They were solubilized in solubilization buffer (50 mM 3-(cyclohexylamino)- 1-propane sulphonic acid (CAPS) and 3% N-lauroylsarcosine) and refolded by dialysis first in 20 mM Tris-HCl (pH 8.5) with 0.1 mM DTT, then in 20 mM Tris-HCl only. The refolded intein-PPSIG proteins were then semi-purified on a preparative SDS-PAGE (see section 2.5) for rabbit immunization. For thio-PPSIG, the proteins were loaded on a nickel column and eluted with 50 mM of imidazole in column wash buffer (20 mM sodium phosphate, 500 mM NaCl, pH 6.0) before it was semi-purified.

2.5 Semi-purification of PPSIG fusion proteins by preparative SDS-PAGE

About 1 mg of refolded fusion proteins were prepared in sample buffer (25 mM Tris-HCl, pH 6.8, 5% glycerol, 1% sodium dodecyl sulphate, 0.01% bromophenol blue and 14.4 mM beta-mercaptoethanol) and boiled for 10 min before loading on an 16 cm x 18 cm x 1.5 mm preparative SDS-PAGE gel [separating gel: 8% or 10 % acrylamide/bis-acrylamide (37.5:1) as indicated in the figures, 0.1 % SDS,

0.375 M Tris-HCl, pH 8.8; stacking gel: 4% acrylamide/bis-acrylamide (37.5:1), 0.1 % SDS, 0.125 M Tris-HCl, pH 6.8]. The gel was assembled in an Amersham SE 600 standard dual cooled vertical unit (Buckinghamshire, England) and electrophoresed at 50 V for 16 h (running buffer: 25 mM Tris, 192 mM glycine, 0.1 % SDS). The gel was stained with 1% Coomassie brilliant blue R in water and destained with water. Protein band was excised out and homogenized in Phosphate Buffered Saline (137 mM sodium chloride, 2.7 mM potassium chloride, 10.1 mM dibasic sodium phosphate and 1.7 mM monobasic potassium phosphate) in a Teflon-glass homogenizer using a Wheaton overhead stirrer.

2.6 Rabbit immunization

The homogenized protein band containing the desired fusion protein antigen was mixed with the same volume of Sigma Freund's adjuvant (St. Louis, MO, USA) and passed through two syringes until emulsification occurred. Pre-immune blood was collected from two 3-month old New Zealand White rabbits from their ear veins. Each fusion protein antigen was injected subcutaneously at the neck of the rabbit following standard protocols (Kakar *et al.* 2001; Marikar *et al.* 2006; Poland *et al.* 1991). Five boosts were injected at 1-month intervals and blood was collected 10-14 days after each boost. Blood collected was allowed to coagulate at room

temperature for 1 h, followed by two consecutive centrifugations at 2,500 g for 20 min at room temperature. The supernatant was stored as serum for Western blotting.

2.7 Northern blotting analysis

2.7.1 Probe preparation

DIG-labelled cDNA probe was produced using the RACE clone (5'II32) amplified with a forward primer 5'-GAGCCCTGGCAGC AACAT-3' and a reverse primer 5'-GCCTTGACCTTTGAGCC-3' (Figure 2.5). The PCR reaction contained 20 ng RACE DNA template, 2 U Taq DNA polymerase (Genesys, UK), 1.5 mM $MgCl_2$, 1 μM forward primer, 1 μM reverse primer, 0.2 mM PCR DIG labelling mix and 1X PCR buffer in 50 μl autoclaved distilled water. The reaction was performed first at 95°C for 7 min, followed by 30 cycles at 95°C for 30 s, 60°C for 1 min and 72°C for 2 min 15 s using a thermocycler (Applied Biosystem GeneAMP PCR 9700).

2.7.2 Formaldehyde-agarose gel electrophoresis, blotting of RNA and hybridization

Northern blotting analysis was carried out using samples from mice described in section 2.2. PPARα wild-type and knockout mice were sacrificed by cervical dislocation and their livers were obtained. About 0.5 g of each liver was

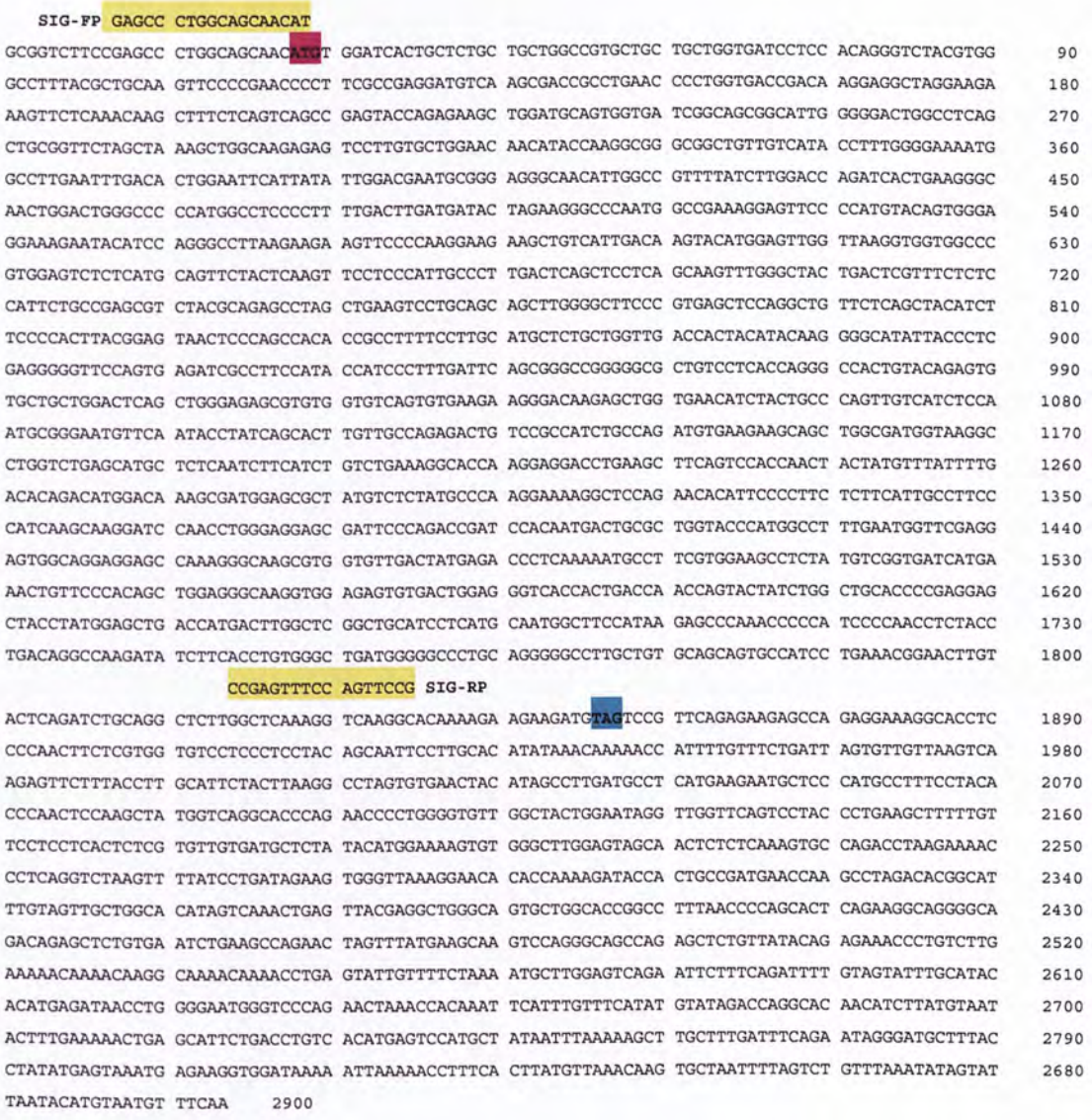


Figure 2.5 Primer design for DIG-labelled cDNA probe preparation for Northern blotting analysis

A set of primers (highlighted in yellow) were designed to amplify a DIG-labelled cDNA probe of ~1800 bp containing the full open reading frame on PPSIG mRNA. The start codon and stop codon of PPSIG are highlighted in blue and red respectively.

homogenized and its total RNA was extracted by TRIzol® reagent using standard protocol provided by the manufacturer. RNA pellet was dissolved in formamide and its concentration was determined by its absorbance at 260 nm. A 30 μg of each sample prepared in core mix (5% formaldehyde, 0.6X loading dye in 1X MOPS) was denatured at 68°C and resolved on a 1%-formaldehyde agarose gel (1% formaldehyde, 1% agarose in 1X MOPS). After electrophoresed at 25 V for 16 h, the gel was incubated in 10X SSC (3 M sodium chloride, 0.3 M sodium citrate, pH 7.0) for 45 min and the sample was transferred to a positively charged N^+ nylon membrane (Pall, East Hills, NY, USA) by upward capillary force overnight. After washing in 5X SSC, the membrane was baked at 80°C for 2 h in an oven (Memmert UM300) in order to fix the RNA onto the membrane. The membrane was pre-hybridized in DIG-Easy hyb buffer (Roche, Germany) with gentle shaking at 42°C for 2 h. Hybridization was carried out overnight with shaking at 42°C in DIG-Easy hyb buffer containing the DIG-labelled cDNA probe described in section 2.7.1. The membrane was then washed first in low stringency buffer (0.1% SDS in 2X SSC) twice for 5 min at room temperature and then in high stringency buffer (0.1%SDS in 0.5X SSC) twice for 5 min at 68°C. After blocking the membrane in 1X blocking buffer [1% (g/v) blocking reagent in 0.1 M maleic acid and 0.15 M sodium chloride, pH 7.5] for 1 h, the membrane was incubated in anti-DIG-AP (Roche, Germany) solution (1:10000

diluted in 1X blocking buffer) for 1 h, before it was washed in washing buffer [0.3% (v/v) Tween 20® in 0.1 M maleic acid and 0.15 M sodium chloride, pH 7.5]. The signal was developed in nitroblue tetrazolium chloride (NBT) and 5-bromo-4-chloro-3-indolyl-phosphate-4-toluidine salt (BCIP) solution (Roche, Mannheim, Germany) and captured with a scanner (Epson GT-9500).

2.8 Subcellular fractionation

Subcellular fractionation was carried out with modifications from several protocols (Bektas *et al.* 2005; Dressman *et al.* 2000; Le Stunff *et al.* 2002; Okado-Matsumoto and Fridovich 2001). About 1 g of liver was minced and added to 30 ml with homogenizing buffer (0.1 M sodium phosphate, pH 7.4, 0.125 M potassium chloride, 1 mM ethylenediaminetetra acetic acid, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM dithiothreitol and 0.25 M sucrose) in a Teflon-glass homogenizer. The livers were homogenized using the Wheaton overhead stirrer with a speed no more than 2. The centrifugation procedure was summarized in Figure 2.6. The homogenate was centrifuged in Himac CR 21 centrifuge using rotor 23 at 1,000 g for 10 min at 4°C to obtain the pellet P1 (nucleus). The resultant supernatant was centrifuged at 3,000 g for 10 min at 4°C to obtain the second pellet P2 (heavy mitochondria). The third pellet P3 (light mitochondria) was

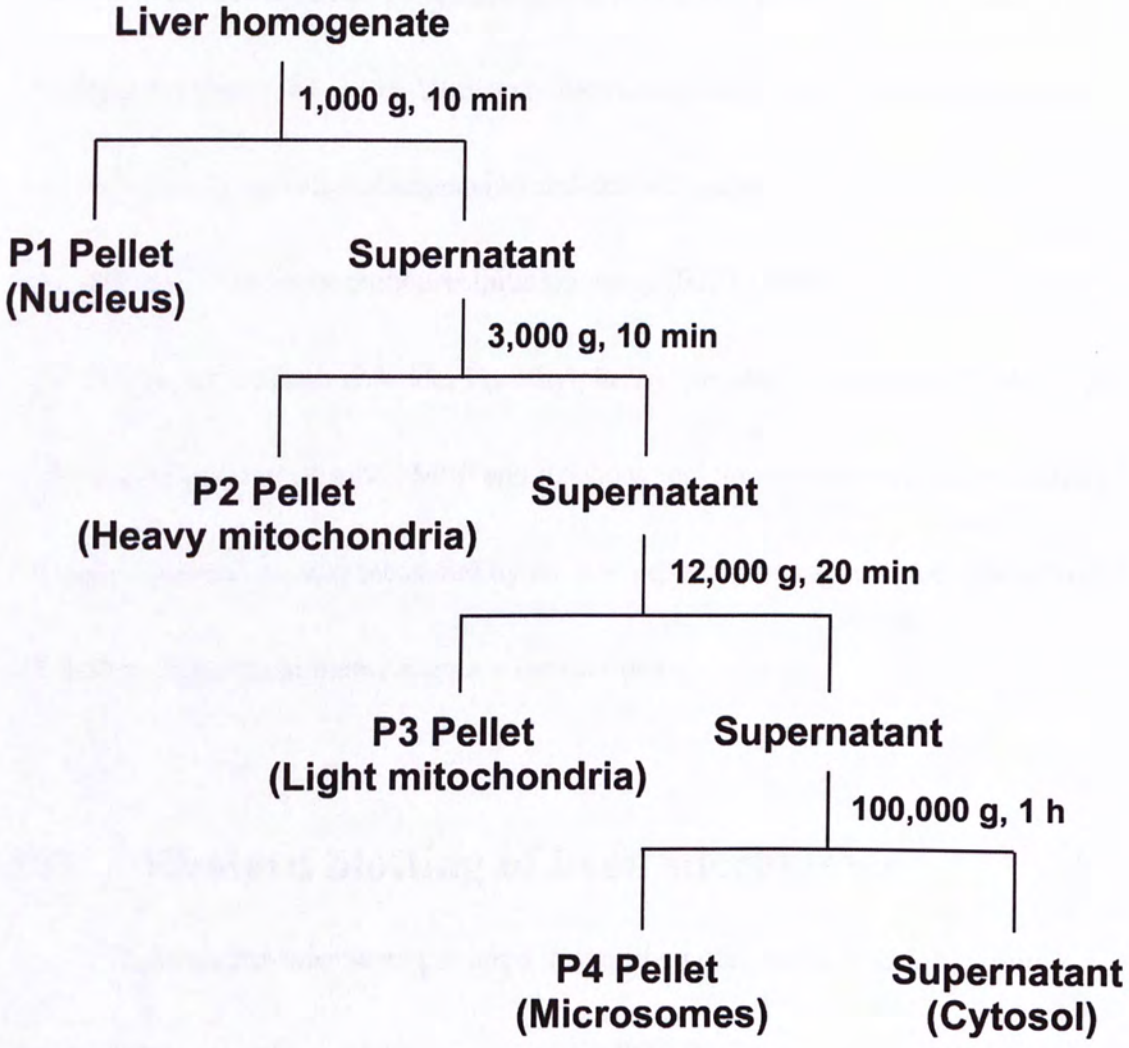


Figure 2.6 Flowchart for subcellular fractionation procedure

Liver homogenate was centrifuged at 1,000 g to pellet nucleus (P1), 3,000 g to pellet heavy mitochondria (P2) and 12,000 g to pellet light mitochondria (P3). Microsomes (P4) were pellet by centrifugation at 100,000 g and the supernatant was referred to as cytosol.

obtained by centrifugation at 12,000 g for 20 min at 4°C. Lastly, a P4 pellet (microsomes) was obtained by centrifugation at 100,000 g for 1 h at 4°C in a Beckman Optima™ XL-100K Ultracentrifuge using rotor 70Ti. Each pellet was washed twice by repeated resuspension and centrifugation. The pellet was finally homogenized in radio-immunoprecipitation assay (RIPA) buffer (50 mM Tris-HCl, pH 7.5, 150 mM sodium chloride, 1% ethylphenyl-polyethylene glycol (NP-40), 0.5% sodium deoxycholate, 1 mM PMSF and Roche complete protease inhibitor cocktail). Protein concentration was measured by BCA™ protein assay kit (Pierce, Rockford, IL, USA) according to the manufacturer's instructions.

2.9 Western blotting of liver microsomes

Liver microsomes were prepared in sample buffer and boiled for 10 min before loading on a 16 cm x 18 cm x 0.75 mm SDS-PAGE gel. The gel was electrophoresed at 50 V for 16 h using an Amersham SE 600 standard dual cooled vertical unit, and proteins were blotted onto a 0.45 μm Pall BioTrace™ PVDF Transfer Membrane (East Hills, NY, USA) at 70 V for 1 h using an Amersham TE 62 transfer cooled unit (Buckinghamshire, England) (transfer buffer: 25 mM Tris, 192 mM glycine, 20% methanol). The blocked membrane was incubated with shaking on an orbital shaker platform (Scienceware, Spindrive®) with primary antibody

solutions overnight at 4°C, then with secondary antibody solution (Zymed alkaline phosphatase (AP)-goat anti-rabbit IgG, 1:2000; Zymed AP-goat anti-mouse IgG, 1:2000) for 2 h. After washing, membrane was developed colourimetrically using NBT and BCIP as substrate (Roche, Mannheim, Germany) and captured with a scanner (Epson GT-9500).

2.10 Immunoprecipitation

Immunoprecipitation was carried out with modifications according to several protocols (Crockett *et al.* 2004; Kakar *et al.* 2001; Wu *et al.* 2003). A 1:500 dilution of pre-immune serum from the rabbit injected with Thio-PPSIG was added to 500 μg microsomes in 1 ml RIPA buffer for 1 h incubation with shaking at 4°C, followed by addition of 50 μl protein A-agarose (Roche, Germany) for a further 3 h incubation. The pre-immune serum was removed together with protein A-agarose by centrifugation at 12,000 g for 20 s. Then a 1:100 dilution of anti-PPSIG antiserum was added to the microsomes for 1 h incubation, followed by addition of 50 μl protein A-agarose for incubation overnight at 4°C. The agarose pellet was collected by centrifugation at 12,000 g for 20 s and washed according to manufacturer's instruction (Roche, Germany). The final pellet was resuspended in 50 μl of sample buffer and boiled for 10 min. After centrifugation at 12,000 g for 20 s, the

supernatant was resolved on an 8-18% gradient SDS-PAGE gel for Western blotting analysis using anti-PPSIG antiserum.

2.11 Mass spectrometry

2.11.1 Trypsin digestion and peptide extraction

Digestion and extraction peptides followed a general procedure modified from several protocols (Crockett *et al.* 2004; Goumon *et al.* 2006; Suzuki *et al.* 2006). Recombinant or immunoprecipitated native PPSIG was resolved on SDS-PAGE gels and the protein band was excised into 5-10 discs having a diameter of 1 mm with a blunt-end needle. The gel discs were washed with 500 µl of 25 mM ammonium bicarbonate (NH_4HCO_3) for 1 h with vortex. They were then dehydrated by three consecutives washes each with 100 µl of 50% acetonitrile / 25 mM NH_4HCO_3 for 10 min, followed by a final wash of 100 µl 100% acetonitrile for 10 min. The gel was flushed to complete dryness using a speed vacuum (Savant SC110A Speed Vac® Plus). Proteins in the gel discs were reduced by 30 µl of 10 mM DTT / 25 mM NH_4HCO_3 at 60°C for 30 min. After discarding the DTT solution, 30 µl of 55 mM iodoacetamide / 25 mM NH_4HCO_3 was added to the gel discs to alkylate the proteins for 30 min in dark. The gel discs were washed and dehydrated to complete dryness as described previously, and rehydrated with 20 µl sequencing grade modified trypsin (Promega,

Madison, WI, USA) solution (20 ng/μl trypsin / 25 mM NH_4HCO_3) for 30 min on ice. Another 10 μl of 25 mM NH_4HCO_3 was added to cover the gel pieces before incubating the gel at 37°C overnight. Digested peptides were extracted with 20 μl of 25 mM NH_4HCO_3, 20 μl of 50% acetonitrile / 5% trifluoroacetic acid (TFA) for 3 times and 20 μl of 100% acetonitrile consecutively in a sonicating water bath. The extracted peptides were flushed to dryness using a speed vacuum, resuspended in 0.1% TFA and purified by passing through $ZipTip_{C18}$ with a column bed of 0.6 μl (Millipore, Bedford, MA, USA) following manufacturer's protocol. Finally, peptides were eluted in 3 μl of 50% acetanitrile / 0.1% TFA.

2.11.2 Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry

After addition of equal amount of matrix (α-cyano-4-hydroxycinnamic acid), 0.5 μl of sample was applied on a MALDI plate for mass spectrometry analysis on the Applied Biosystems 4700 Proteomics Analyzer. Peptide mass fingerprinting (PMF) spectrum of each peptide spot was generated automatically by plotting the relative intensity of positively charged ions against mass-to-charge (*m/z*) ratio of the ions in the range of 500-5000. High intensity peptide peaks were sequenced by MS/MS by generation of fragment ion spectra with an MS fragment tolerance of 0.5 Da. Data

from PMF were processed by the Mascot software using the GPS Explorer™ Workstation. After removing peaks resulting from trypsin autolysis, the peptide peaks were submitted for searching with in the NCBI non-redundant database with no missed cleavage allowed and precursor mass tolerance of 100 ppm.

Chapter 3 Results

3.1 Cloning of PPSIG into pThioHis and pTYB vectors

3.1.1 Cloning of PPSIG into pThioHis vector

A ~800 bp partial length of PPSIG cDNA insert containing *Sac* II and *Xba* I digestion sites was amplified (Figure 3.1). Ligation of *Sac* II and *Xba* I digested PPSIG insert with a ~4.4 kb *Sac* II and *Xba* I digested pThioHis C vector (Figure 3.2) produced a pThioHis-PPSIG construct of ~5.2 kb. Screening of recombinant clones by phenol/chloroform method identified 13 clones containing plasmid having the same size as pThioHis-PPSIG construct (Figure 3.3A). To confirm the identity of the recombinants, 4 clones (3, 8, 9 and 12) were selected for DNA plasmid extraction followed by restriction enzyme digestion. Double digestion with *Sac* II and *Xba* I yielded 2 bands of ~4.4 kb and ~0.8 kb (Figure 3.3B). After DNA sequencing, clone 9 was found to be 99% identical to the 5'II32 RACE clone generated from the FDD fragment with a mutation from leucine to proline (Figure 3.4). Its DNA plasmid was then transformed into BL21 star (DE3) competent cells for protein expression.

3.1.2 Cloning of PPSIG into pTYB vector

The same partial length of PPSIG cDNA insert of ~800 kb containing *Nde* I

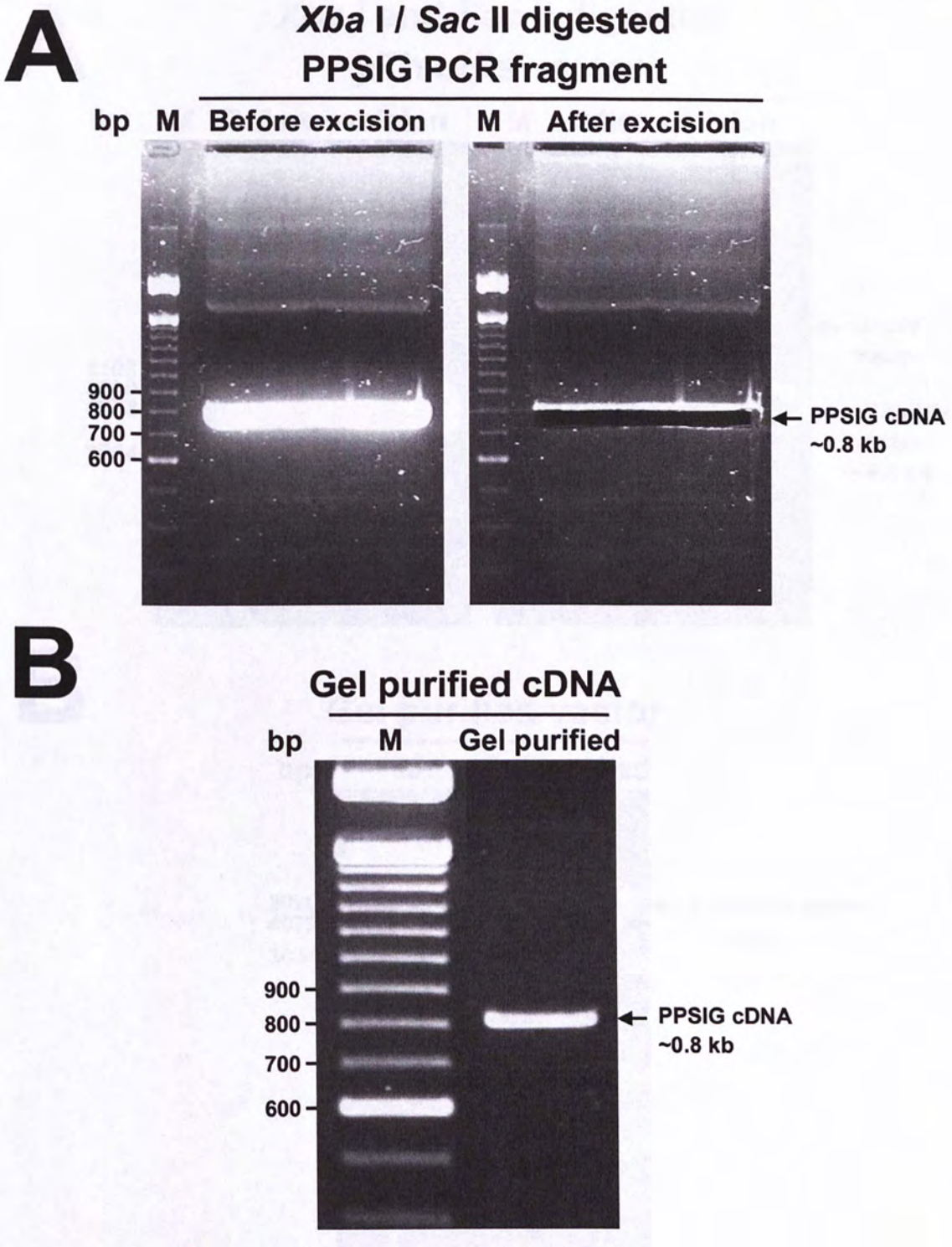


Figure 3.1 Preparation of *Xba* I and *Sac* II digested partial PPSIG cDNA fragment for cloning into pThioHis vector

(A) Electrophoresis of PCR amplified and *Xba* I and *Sac* II digested PPSIG cDNA fragment on a 1.5 % agarose gel for gel excision and purification. (B) Gel purified PPSIG cDNA insert. M, 100 bp DNA ladder.

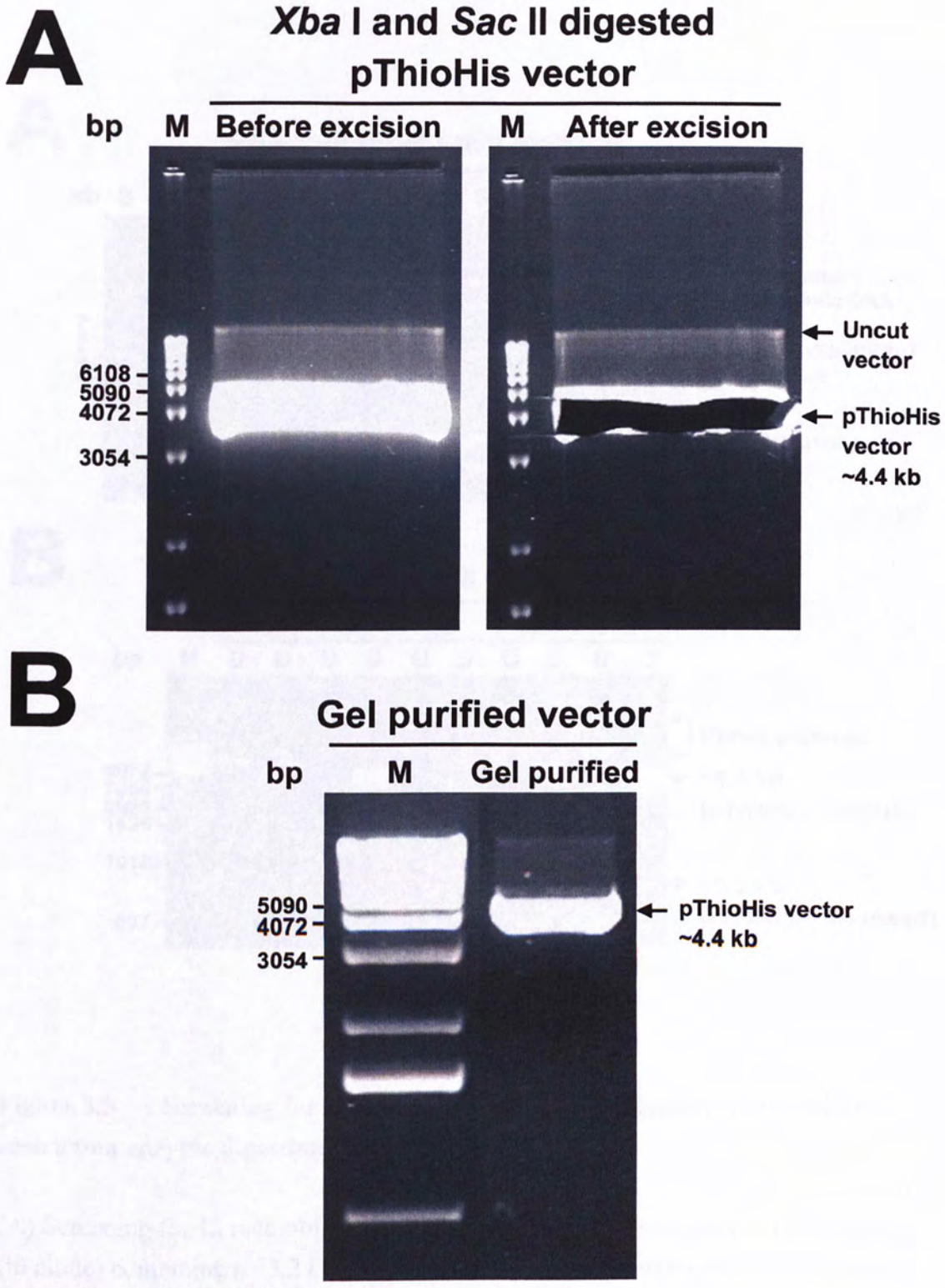


Figure 3.2 Preparation of *Xba* I and *Sac* II digested pThioHis vector

(A) Electrophoresis of *Xba* I and *Sac* II digested pThioHis vector on a 1% agarose gel for gel excision and purification. (B) Gel purified pThioHis vector. M, 1 kb DNA ladder.

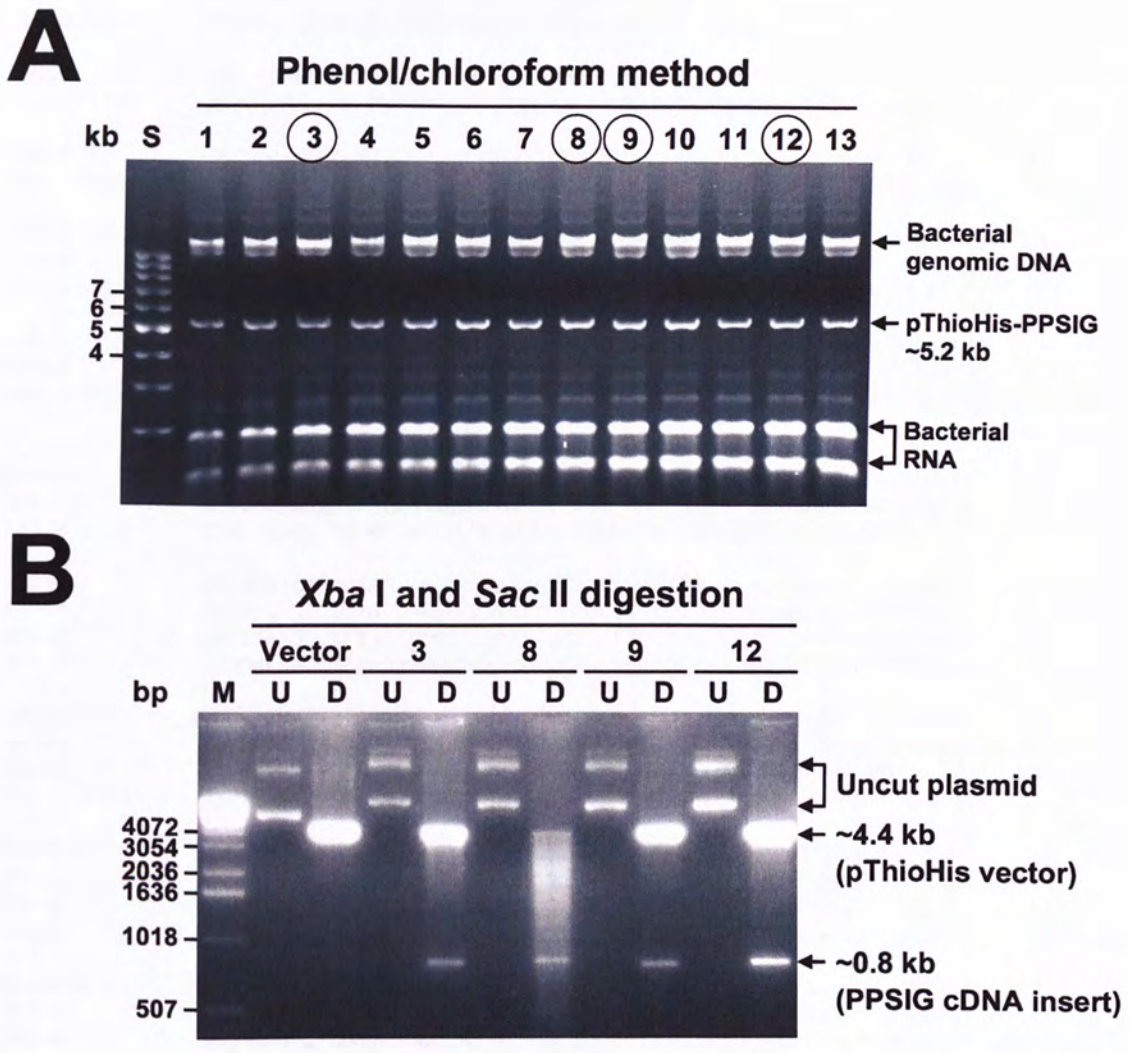


Figure 3.3 Screening for recombinants by phenol/chloroform method and restriction enzyme digestion

(A) Screening for 13 recombinants using phenol/chloroform method. Four clones (in circle) containing a ~5.2 kb pThioHis-PPSIG construct were selected for DNA plasmid extraction. (B) Restriction enzyme digestion of the extracted plasmid DNA with *Xba* I and *Sac* II. U, uncut plasmid DNA. D, *Xba* I and *Sac* II digested plasmid DNA. S, supercoiled DNA ladder. M, 1 kb DNA ladder.

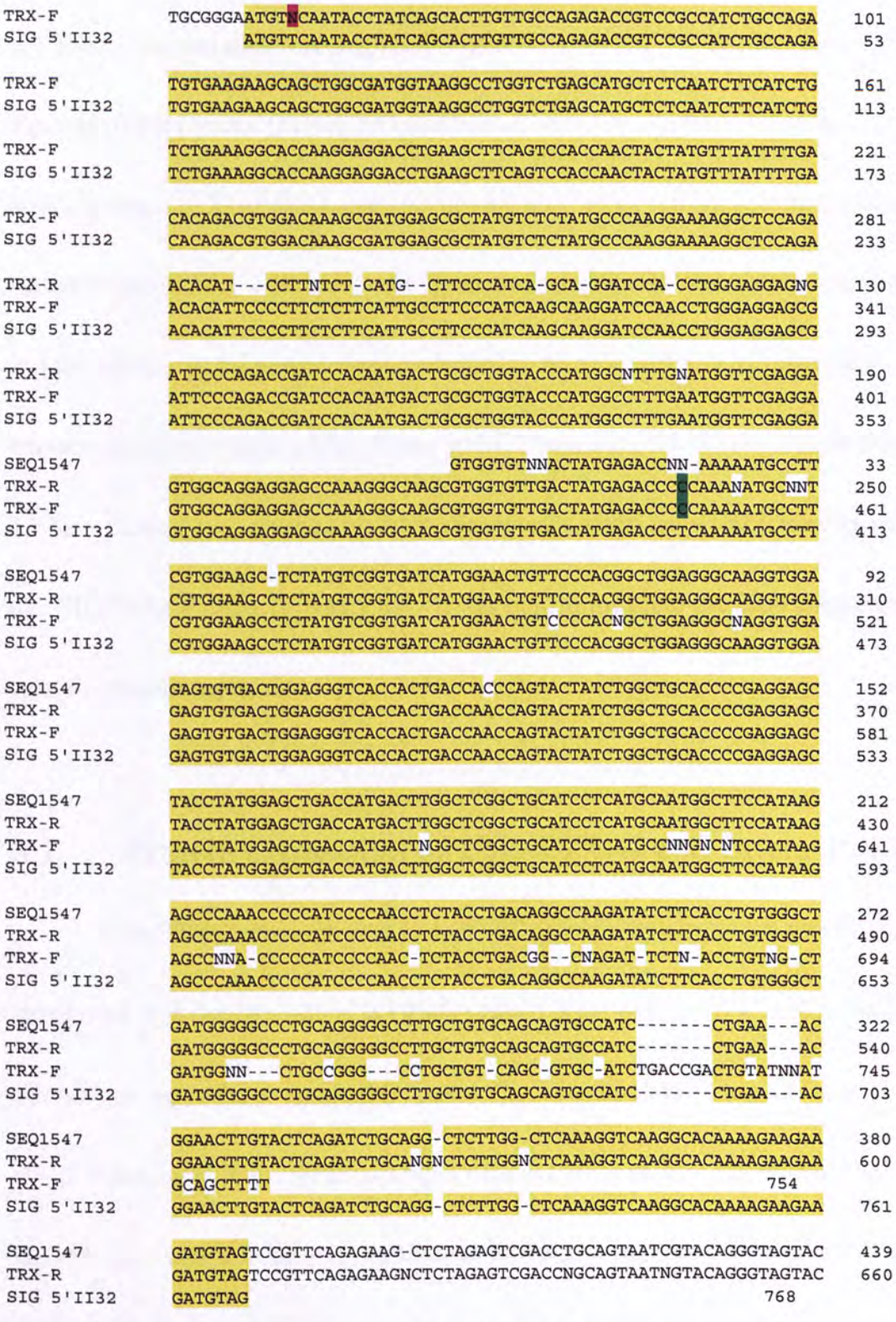


Figure 3.4 DNA sequencing of pThioHis-PPSIG plasmid

The DNA sequence of clone 9 was 99% identical to the RACE clone sequence (SIG 5'II32), with a mismatch at the 5' position (highlighted in purple) and a mutation from T (401) to C (highlighted in green) that changes the resulting amino acid from leucine to proline. It was used for expression of Thio-PPSIG fusion protein.

and *Sap* I digestion sites was amplified (Figure 3.5). The ~7.5 kb *Nde* I and *Sap* I digested pTYB1 vector (Figure 3.6) was ligated with *Nde* I and *Sap* I digested PPSIG insert, giving a pTYB-PPSIG construct of ~8.3 kb. After screening of recombinant clones (Figure 3.7A), 2 clones (5 and 9) were selected for DNA plasmid extraction and the identity of the plasmid was confirmed to be pTYB-PPSIG by restriction enzyme digestion with *EcoR* V, which yielded 2 bands at ~5.5 kb and ~2.8 kb (Figure 3.7B). Clone 5 was selected for DNA sequencing, which gave 100% identity with the 5'II32 RACE clone (Figure 3.8). It was transformed into ER 2566 competent cells for protein expression.

3.2 Protein expression of Thio-PPSIG and Intein-PPSIG

Thio-PPSIG was expressed as a ~43 kDa fusion protein, which was in agreement with that of a ~28 kDa PPSIG protein fused with a ~15 kDa His-Patch thioredoxin tag at the N-terminal (Figure 3.9). The thioredoxin tag was expressed as a ~15 kDa protein in *E. coli* transformed with pThioHis vector. Both could be lighted up by anti-thio antibody in Western blottings. Similarly, Intein-PPSIG was expressed as a ~87 kDa protein, which corresponded to a ~28 kDa PPSIG fused with a ~59 kDa CBD-intein tag at the C-terminal (Figure 3.10). Both intein-PPSIG and the intein tag were lighted up by anti-CBD antiserum, confirming the presence of

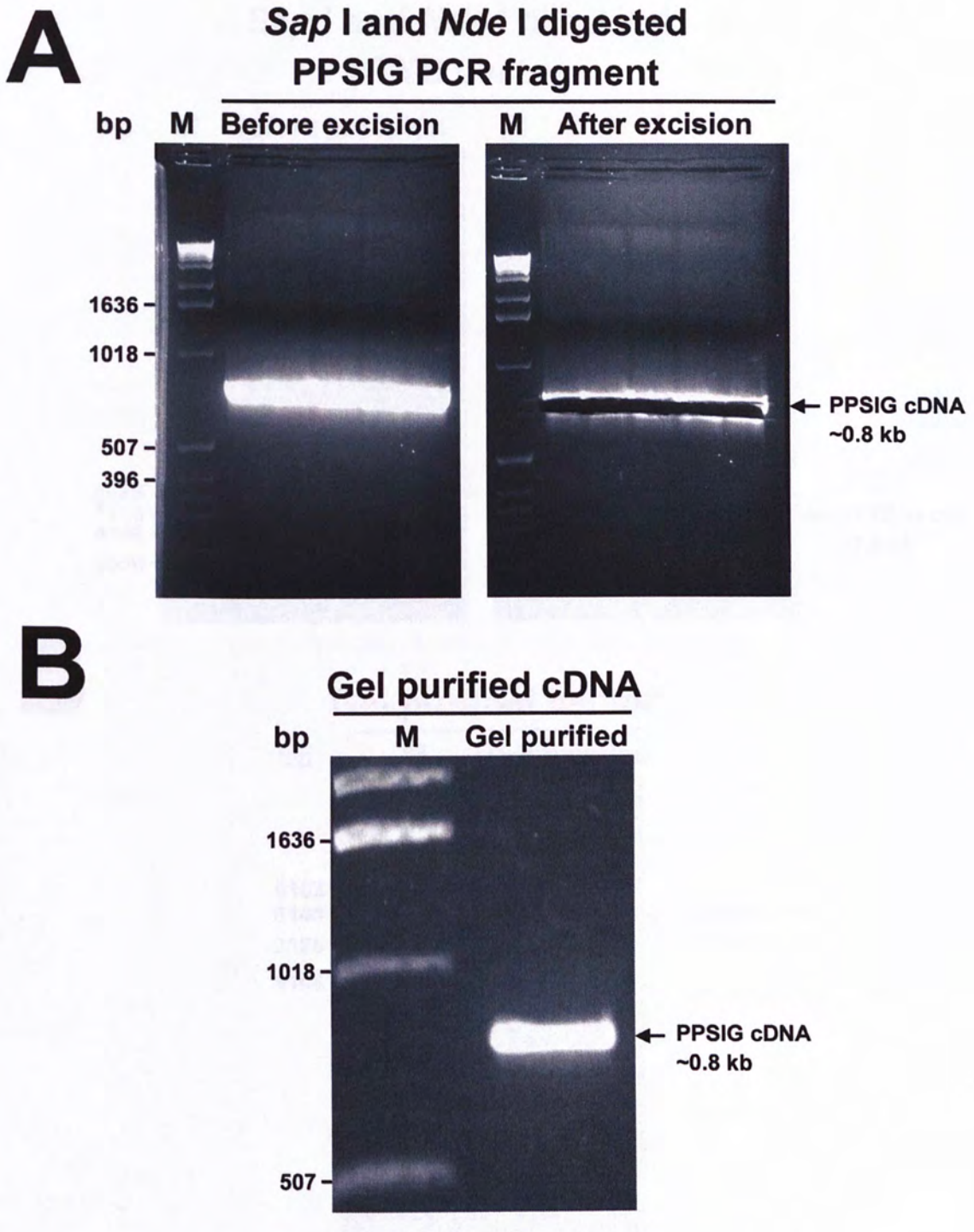


Figure 3.5 Preparation of *Sap* I and *Nde* I digested partial PPSIG cDNA fragment for cloning into pTYB1 vector

(A) Electrophoresis of PCR amplified *Sap* I and *Nde* I digested PPSIG cDNA fragment on a 1.5 % agarose gel for gel excision and purification. (B) Gel purified PPSIG cDNA insert. M, 1 kb DNA ladder.

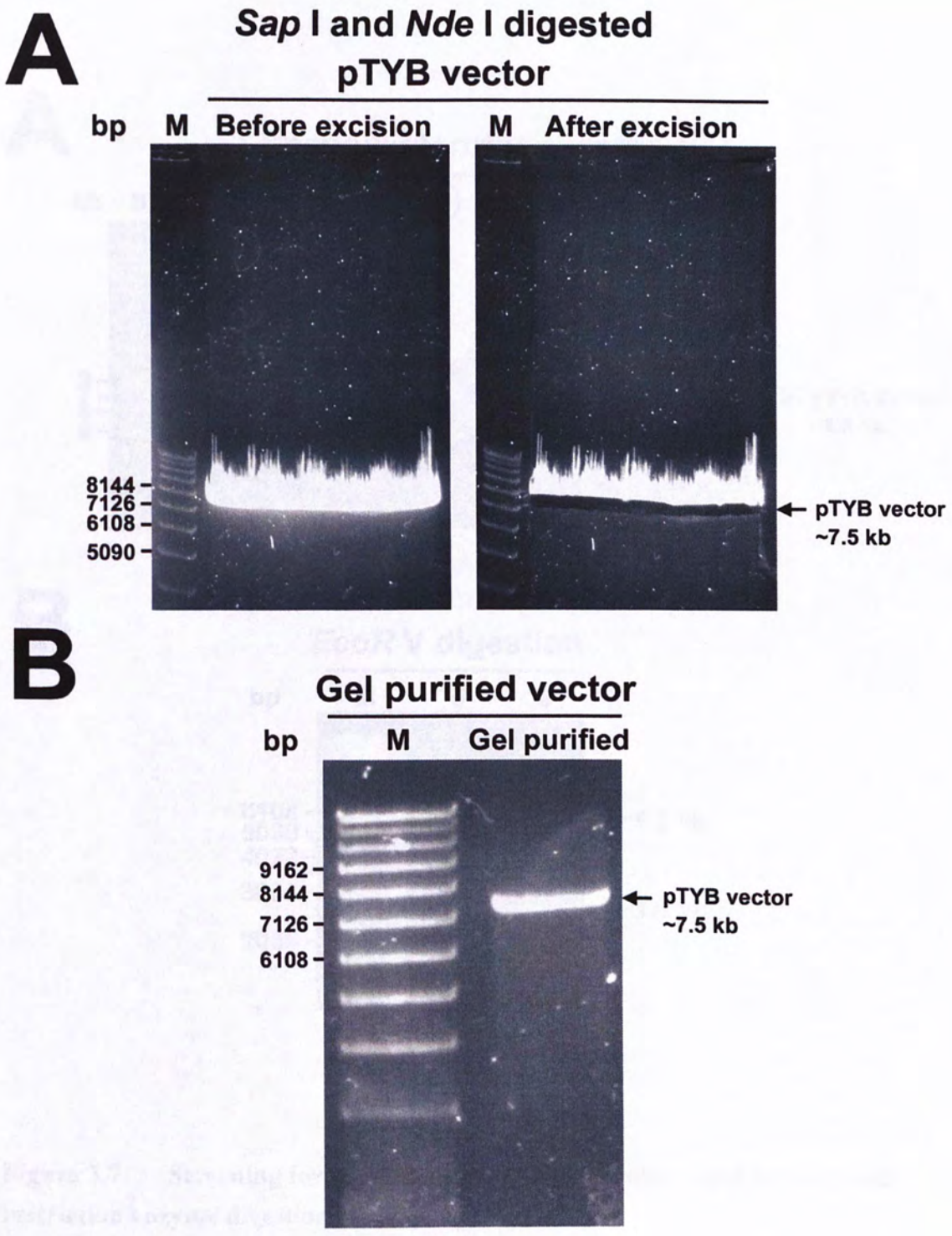


Figure 3.6 Preparation of *Sap* I and *Nde* I digested pTYB1 vector

(A) Electrophoresis of *Sap* I and *Nde* I digested pTYB vector on a 0.6% agarose gel for gel excision and purification. (B) Gel purified pTYB vector. M, 1 kb DNA ladder.

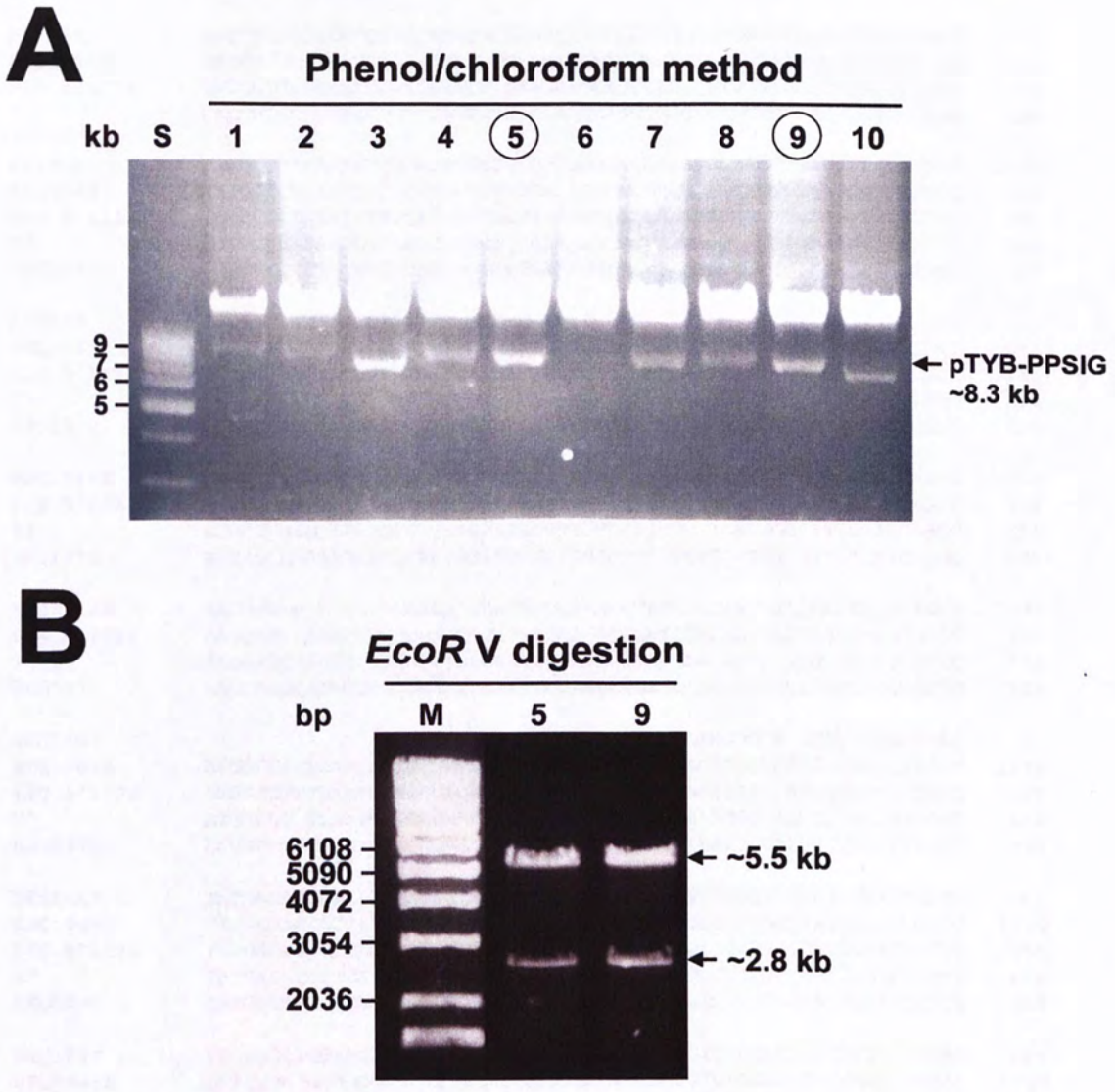


Figure 3.7 Screening for recombinants by phenol/chloroform method and restriction enzyme digestion

(A) Screening for 10 recombinant clones using phenol/chloroform method. Two clones (in circle) containing a ~8.3 kb pTYB-PPSIG construct were chosen for DNA plasmid extraction. (B) Restriction enzyme digestion of the extracted plasmid DNA with *EcoR* V. S, supercoiled DNA ladder. M, 1 kb DNA ladder.

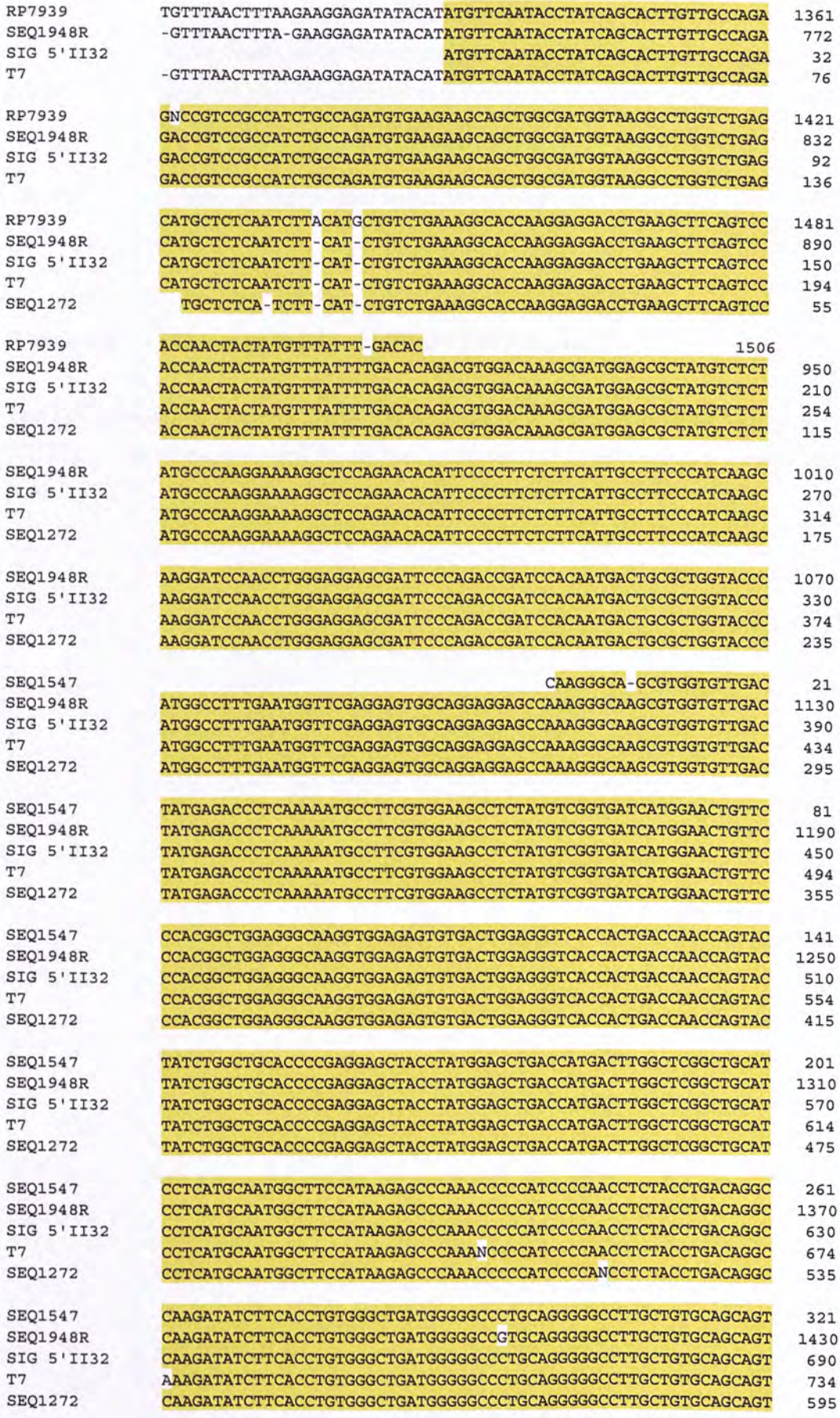

RP7939 TGTTTAACTTTAAGAAGGAGATATACATATGTTCAATACCTATCAGCACTTGTTGCCAGA 1361
SEQ1948R -GTTTAACTTTA-GAAGGAGATATACATATGTTCAATACCTATCAGCACTTGTTGCCAGA 772
SIG 5'II32 ATGTTCAATACCTATCAGCACTTGTTGCCAGA 32
T7 -GTTTAACTTTAAGAAGGAGATATACATATGTTCAATACCTATCAGCACTTGTTGCCAGA 76
RP7939 GNCCGTCCGCCATCTGCCAGATGTGAAGAAGCAGCTGGCGATGGTAAGGCCTGGTCTGAG 1421
SEQ1948R GACCGTCCGCCATCTGCCAGATGTGAAGAAGCAGCTGGCGATGGTAAGGCCTGGTCTGAG 832
SIG 5'II32 GACCGTCCGCCATCTGCCAGATGTGAAGAAGCAGCTGGCGATGGTAAGGCCTGGTCTGAG 92
T7 GACCGTCCGCCATCTGCCAGATGTGAAGAAGCAGCTGGCGATGGTAAGGCCTGGTCTGAG 136
RP7939 CATGCTCTCAATCTTACATGCTGTCTGAAAGGCACCAAGGAGGACCTGAAGCTTCAGTCC 1481
SEQ1948R CATGCTCTCAATCTT-CAT-CTGTCTGAAAGGCACCAAGGAGGACCTGAAGCTTCAGTCC 890
SIG 5'II32 CATGCTCTCAATCTT-CAT-CTGTCTGAAAGGCACCAAGGAGGACCTGAAGCTTCAGTCC 150
T7 CATGCTCTCAATCTT-CAT-CTGTCTGAAAGGCACCAAGGAGGACCTGAAGCTTCAGTCC 194
SEQ1272 TGCTCTCA-TCTT-CAT-CTGTCTGAAAGGCACCAAGGAGGACCTGAAGCTTCAGTCC 55
RP7939 ACCAACTACTATGTTTATTT-GACAC 1506
SEQ1948R ACCAACTACTATGTTTATTTTGACACAGACGTGGACAAAGCGATGGAGCGCTATGTCTCT 950
SIG 5'II32 ACCAACTACTATGTTTATTTTGACACAGACGTGGACAAAGCGATGGAGCGCTATGTCTCT 210
T7 ACCAACTACTATGTTTATTTTGACACAGACGTGGACAAAGCGATGGAGCGCTATGTCTCT 254
SEQ1272 ACCAACTACTATGTTTATTTTGACACAGACGTGGACAAAGCGATGGAGCGCTATGTCTCT 115
SEQ1948R ATGCCCAAGGAAAAGGCTCCAGAACACATTCCCCTTCTCTTCATTGCCTTCCCATCAAGC 1010
SIG 5'II32 ATGCCCAAGGAAAAGGCTCCAGAACACATTCCCCTTCTCTTCATTGCCTTCCCATCAAGC 270
T7 ATGCCCAAGGAAAAGGCTCCAGAACACATTCCCCTTCTCTTCATTGCCTTCCCATCAAGC 314
SEQ1272 ATGCCCAAGGAAAAGGCTCCAGAACACATTCCCCTTCTCTTCATTGCCTTCCCATCAAGC 175
SEQ1948R AAGGATCCAACCTGGGAGGAGCGATTCCCAGACCGATCCACAATGACTGCGCTGGTACCC 1070
SIG 5'II32 AAGGATCCAACCTGGGAGGAGCGATTCCCAGACCGATCCACAATGACTGCGCTGGTACCC 330
T7 AAGGATCCAACCTGGGAGGAGCGATTCCCAGACCGATCCACAATGACTGCGCTGGTACCC 374
SEQ1272 AAGGATCCAACCTGGGAGGAGCGATTCCCAGACCGATCCACAATGACTGCGCTGGTACCC 235
SEQ1547 CAAGGGCA-GCGTGGTGTTGAC 21
SEQ1948R ATGGCCTTTGAATGGTTCGAGGAGTGGCAGGAGGAGCCAAAGGGCAAGCGTGGTGTTGAC 1130
SIG 5'II32 ATGGCCTTTGAATGGTTCGAGGAGTGGCAGGAGGAGCCAAAGGGCAAGCGTGGTGTTGAC 390
T7 ATGGCCTTTGAATGGTTCGAGGAGTGGCAGGAGGAGCCAAAGGGCAAGCGTGGTGTTGAC 434
SEQ1272 ATGGCCTTTGAATGGTTCGAGGAGTGGCAGGAGGAGCCAAAGGGCAAGCGTGGTGTTGAC 295
SEQ1547 TATGAGACCCTCAAAAATGCCTTCGTGGAAGCCTCTATGTCGGTGATCATGGAACTGTTC 81
SEQ1948R TATGAGACCCTCAAAAATGCCTTCGTGGAAGCCTCTATGTCGGTGATCATGGAACTGTTC 1190
SIG 5'II32 TATGAGACCCTCAAAAATGCCTTCGTGGAAGCCTCTATGTCGGTGATCATGGAACTGTTC 450
T7 TATGAGACCCTCAAAAATGCCTTCGTGGAAGCCTCTATGTCGGTGATCATGGAACTGTTC 494
SEQ1272 TATGAGACCCTCAAAAATGCCTTCGTGGAAGCCTCTATGTCGGTGATCATGGAACTGTTC 355
SEQ1547 CCACGGCTGGAGGGCAAGGTGGAGAGTGTGACTGGAGGGTCACCACTGACCAACCAGTAC 141
SEQ1948R CCACGGCTGGAGGGCAAGGTGGAGAGTGTGACTGGAGGGTCACCACTGACCAACCAGTAC 1250
SIG 5'II32 CCACGGCTGGAGGGCAAGGTGGAGAGTGTGACTGGAGGGTCACCACTGACCAACCAGTAC 510
T7 CCACGGCTGGAGGGCAAGGTGGAGAGTGTGACTGGAGGGTCACCACTGACCAACCAGTAC 554
SEQ1272 CCACGGCTGGAGGGCAAGGTGGAGAGTGTGACTGGAGGGTCACCACTGACCAACCAGTAC 415
SEQ1547 TATCTGGCTGCACCCCGAGGAGCTACCTATGGAGCTGACCATGACTTGGCTCGGCTGCAT 201
SEQ1948R TATCTGGCTGCACCCCGAGGAGCTACCTATGGAGCTGACCATGACTTGGCTCGGCTGCAT 1310
SIG 5'II32 TATCTGGCTGCACCCCGAGGAGCTACCTATGGAGCTGACCATGACTTGGCTCGGCTGCAT 570
T7 TATCTGGCTGCACCCCGAGGAGCTACCTATGGAGCTGACCATGACTTGGCTCGGCTGCAT 614
SEQ1272 TATCTGGCTGCACCCCGAGGAGCTACCTATGGAGCTGACCATGACTTGGCTCGGCTGCAT 475
SEQ1547 CCTCATGCAATGGCTTCCATAAGAGCCCAAACCCCCATCCCCAACCTCTACCTGACAGGC 261
SEQ1948R CCTCATGCAATGGCTTCCATAAGAGCCCAAACCCCCATCCCCAACCTCTACCTGACAGGC 1370
SIG 5'II32 CCTCATGCAATGGCTTCCATAAGAGCCCAAACCCCCATCCCCAACCTCTACCTGACAGGC 630
T7 CCTCATGCAATGGCTTCCATAAGAGCCCAAANCCCCATCCCCAACCTCTACCTGACAGGC 674
SEQ1272 CCTCATGCAATGGCTTCCATAAGAGCCCAAACCCCCATCCCCANCCTCTACCTGACAGGC 535
SEQ1547 CAAGATATCTTCACCTGTGGGCTGATGGGGGCCCTGCAGGGGGCCTTGCTGTGCAGCAGT 321
SEQ1948R CAAGATATCTTCACCTGTGGGCTGATGGGGGCCGTGCAGGGGGCCTTGCTGTGCAGCAGT 1430
SIG 5'II32 CAAGATATCTTCACCTGTGGGCTGATGGGGGCCCTGCAGGGGGCCTTGCTGTGCAGCAGT 690
T7 AAAGATATCTTCACCTGTGGGCTGATGGGGGCCCTGCAGGGGGCCTTGCTGTGCAGCAGT 734
SEQ1272 CAAGATATCTTCACCTGTGGGCTGATGGGGGCCCTGCAGGGGGCCTTGCTGTGCAGCAGT 595

Figure 3.8 DNA sequencing of pTYB-PPSIG plasmid

The DNA sequence of clone 5 was 100% identical to the RACE clone sequence (SIG 5'II32). The clone was used for protein expression to produce Intein-PPSIG fusion protein.

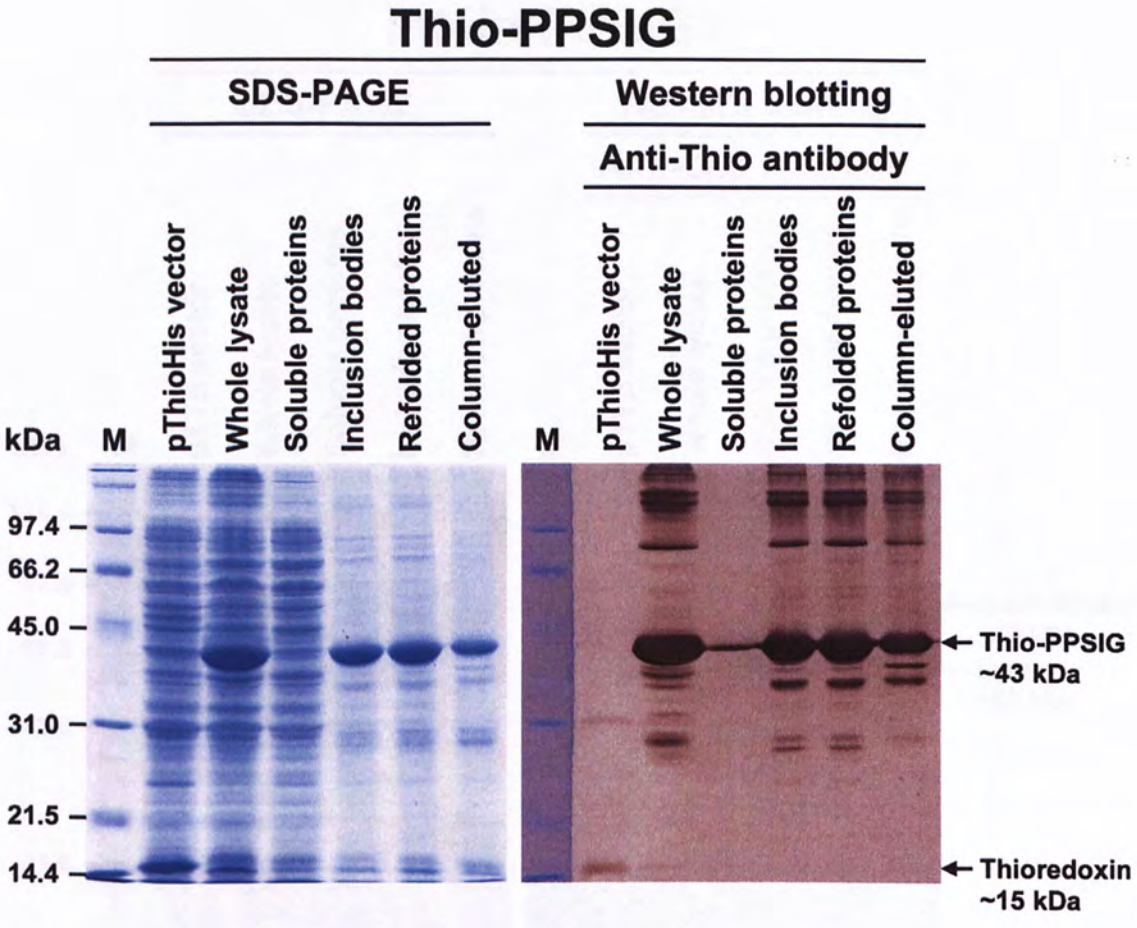


Figure 3.9 Expression and refolding of Thio-PPSIG

Thio-PPSIG was expressed in BL21 star (DE3) cells as a ~43 kDa fusion protein as shown in the SDS-PAGE and was absent in cells transformed only with pThioHis vector, in which a ~15 kDa thioredoxin was expressed. After sonication of bacterial culture, the whole lysate was separated by centrifugation into soluble proteins and inclusion bodies that contains Thio-PPSIG, which was further refolded and eluted with imidazole on a nickel column. Western blotting analysis showed that both Thio-PPSIG and thioredoxin were lighted up by anti-thio monoclonal antibody (1:5000) in all fractions mentioned above. M, unstained protein standards.

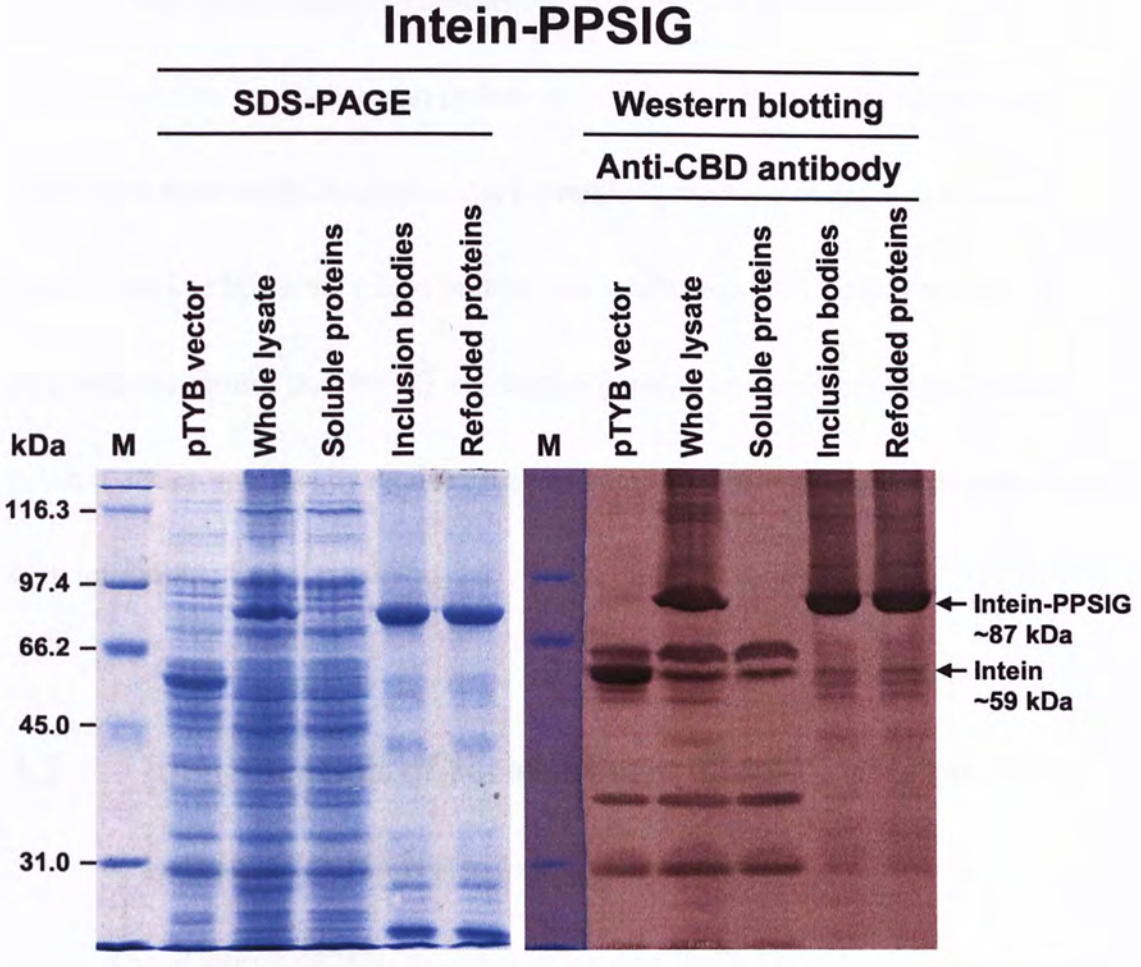


Figure 3.10 Expression and refolding of Intein-PPSIG

Intein-PPSIG was expressed in ER2566 cells as a ~87 kDa fusion protein. In cells transformed only with pTYB vector, a ~59 kDa intein protein was expressed. Whole lysate was separated by centrifugation into soluble proteins and inclusion bodies, which contained Intein-PPSIG and was further refolded to restore its native conformation. Both Intein-PPSIG and intein were lighted up in the Western blotting analysis using 1:5000 anti-(chitin-binding domain) antiserum (anti-CBD) as primary antibody. M, unstained protein standards.

fusion partner in both fusion proteins. Centrifugation of the Thio-PPSIG and Intein-PPSIG whole lysate revealed that the recombinant fusion proteins was insoluble and were pelleted with inclusion bodies, implying that the fusion proteins might have been misfolded during the expression process. In order to restore the native form of PPSIG, the fusion protein was solubilized and refolded by dialysis. The refolded soluble thio-PPSIG was further purified by passing through a nickel column, which specifically retained the His-Patch thioredoxin tag, followed by elution with imidazole.

3.3 Identification of recombinant Thio-PPSIG and Intein-PPSIG by mass spectrometry

The identity of PPSIG fusion proteins were further investigated by its mass spectrum after trypsin digestion. Searching with the peptide masses obtained in the NCBI database identified the protein as the 0610039N19Rik protein (accession number: 15029943) which had the same sequence with the partial PPSIG used for cloning. Peptides recovered matched with the characteristic mass spectrum of trypsin-digested PPSIG. Thio-PPSIG was identified as PPSIG with a score of 240 (significant threshold=78). Peptide fingerprinting results yielded 4 peptide peaks that matched with the trypsin digestion spectrum of the partial length of PPSIG (Table

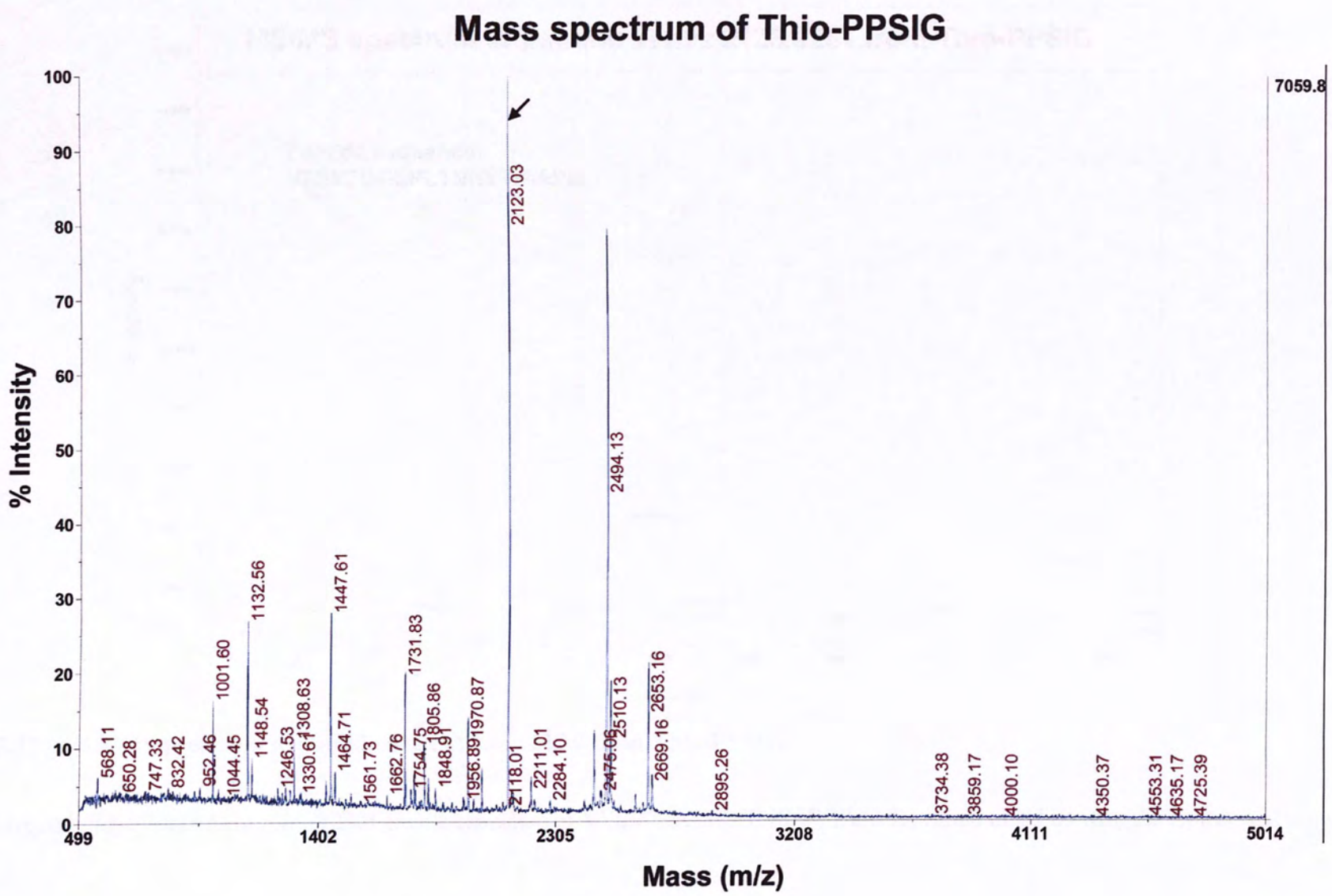
3.1) with sequence coverage of 22.7%. Among the 4 peaks, the peak at *m/z* 2123.0342 indicated on the spectrum (Figure 3.11) had intensity high enough for MS/MS sequencing, which gave a sequence of VESVTGGSPLTNKYYIAAPR (Figure 3.12) located at 157-176 region of the partial PPSIG amino acid sequence (Figure 3.13). When compared with the theoretical sequence VESVTGGSPLTNQYYLAAPR, changes in 2 amino acids were found (Q to K and L to I) and it was due to the coincidence of the fragment ion mass of the amino acid pairs. Similarly, Intein-PPSIG was also identified to be PPSIG with a score of 283 (significant threshold=78) with 4 matching peptide peaks (Table 3.2) having a sequence coverage of 23.5%. The peak with mass 2123.0281 indicated on the mass spectrum (Figure 3.14) gave an MS/MS sequencing result of EVSVTGGSPLTNKYYLAAPR (Figure 3.15), which was located at the same position of the partial PPSIG (Figure 3.16). The change from Q to K was due to the same reason described above. There is also a change in amino acid order of V and E, which may be due to error during ion fragmentation of the first 2 amino acids at the N-terminal of the peptide. After confirming the identity of Thio-PPSIG and Intein-PPSIG, the fusion proteins were used for rabbit immunization for the generation of polyclonal antisera.

Table 3.1 Four peptide peaks from trypsin digestion of Thio-PPSIG identified to match with 0610039N19Rik protein using MALDI-TOF peptide mass fingerprinting

Calculated mass	Observed mass	Position	Amino acid sequence
1132.6045	1132.5638	189-198	LHPHAMASIR
1308.6794	1308.6318	236-247	NLYSDLQALGSK
1766.984	1766.9242	76-91	APEHIPLLFIAFPSSK
2123.0767	2123.0342	157-176	VESVTGGSPLTNQYYLAAPR

Figure 3.11 Mass spectrum of trypsin digested Thio-PPSIG

The masses of peptide peaks from the trypsin digestion spectrum of Thio-PPSIG were submitted for searching in the NCBI database and were found to match with that of the0610039N19Rik protein with a protein score of 240 (significance threshold=78) and a sequence coverage of 22.7%. The peak at *m/z* 2123.0 (arrow head) was analyzed by MS/MS sequencing.


Mass spectrum of Thio-PPSIG
% Intensity
Mass (m/z)
100
90
80
70
60
50
40
30
20
10
0
499
1402
2305
3208
4111
5014
7059.8
2123.03
2494.13
1132.56
1447.61
1001.60
1731.83
1308.63
1148.54
1805.86
1970.87
1464.71
1246.53
1330.61
1662.76
1754.75
1848.91
1956.89
2118.01
2211.01
2284.10
2475.06
2510.13
2653.16
2669.16
2895.26
3734.38
3859.17
4000.10
4350.37
4553.31
4635.17
4725.39
568.11
650.28
747.33
832.42
952.46
1044.45

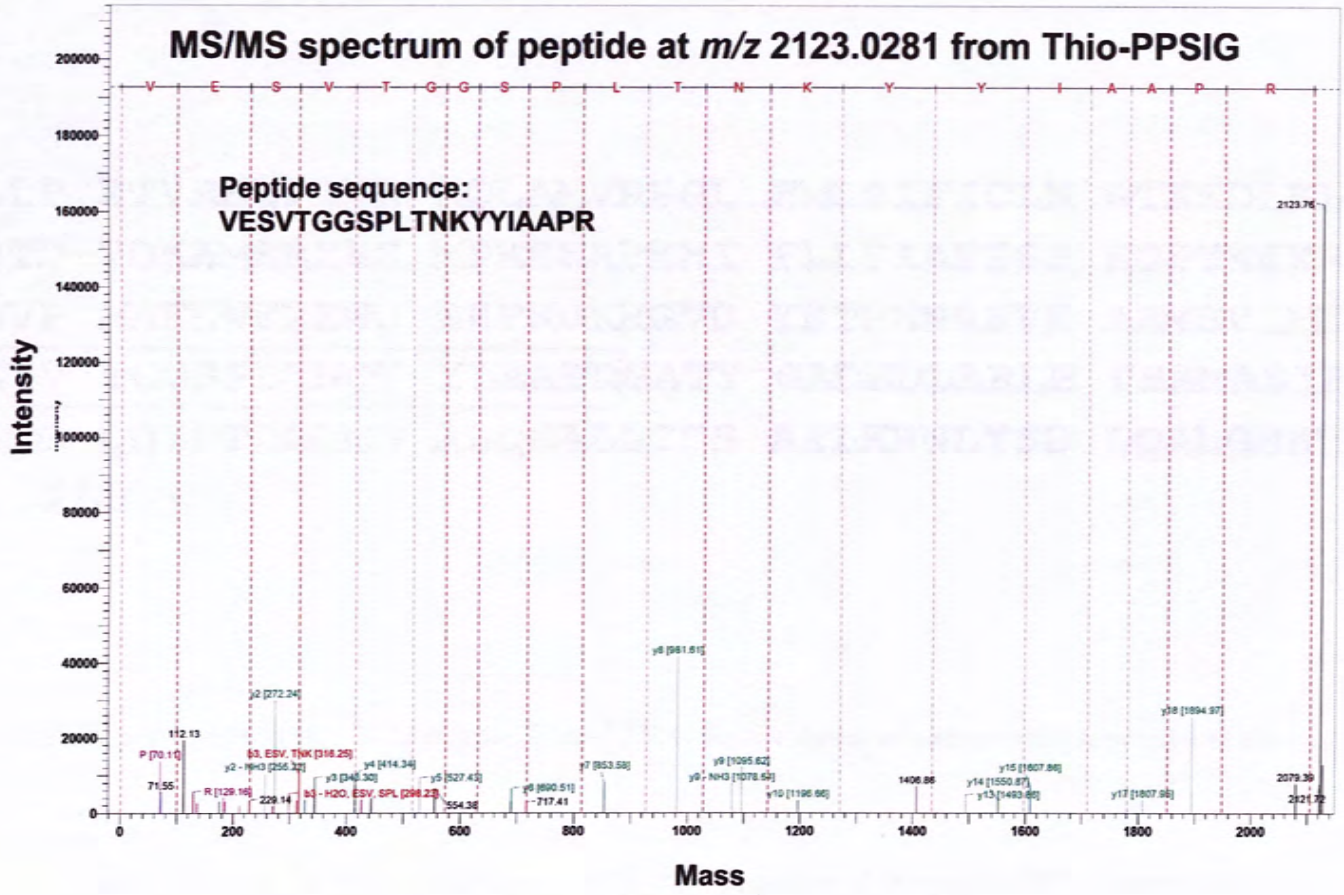


Figure 3.12 MS/MS spectrum of peptide peak at *m/z* 2123.0 from Thio-PPSIG

MS/MS sequencing of peptide at m/z 2123.0 gave a sequence of VESVTGGSPLTNKYYIAAPR, which matched with the amino acid sequence of PPSIG.

MFNTYQHLLP ETVRHLPDVK KQLAMVRPGL SMLSIFICLK GTKEDLKLQS 50
TNYYVYFDTD VDKAMERYVS MPKEKAPEHI PLLFIAFPSS KDPTWEERFP 100
DRSTMTALVP MAFEWFEEWQ EEPKGKRGVD YETPKNAFVE ASMSVIMELF 150
PRLEGKVESV TGGSPLTNQY YLAAPRGATY GADHDLARLH PHAMASIRAQ 200
TPIPNLYLTG QDIFTCGLMG ALQGALLCSS AILKRNLYSD LQALGSKVKA 250
QKKKM 255

Figure 3.13 Position of the peptide at m/z 2123.0 from Thio-PPSIG on the deduced amino acid sequence of the partial PPSIG for cloning

The peptide at m/z 2123.0 was located at the 157-176 region of the protein sequence of the partial PPSIG used in cloning of Thio-PPSIG as shown.

Table 3.2 Four peptide peaks from trypsin digestion of Intein-PPSIG identified to match with 0610039N19Rik protein using MALDI-TOF peptide mass fingerprinting

Calculated mass	Observed mass	Position	Amino acid sequence
1132.6045	1132.5599	189-198	LHPHAMASIR
1748.8789	1748.8330	1-14	MFNTYQHLLPETVR
1766.984	1766.908	76-91	APEHIPLLFIAFPSSK
2123.0767	2123.0281	157-176	VESVTGGSPLTNQYYLAAPR

Figure 3.14 Mass spectrum of trypsin digested Intein-PPSIG

Masses of peptide peaks from the trypsin digestion spectrum of Intein-PPSIG submitted for searching in the NCBI database were found to match with that of the hypothetical protein MMT-7 with a protein score of 283 (significance threshold=78) and a sequence coverage of 23.5%. The peak at *m/z* 2123.0 (arrow head) was analyzed by MS/MS sequencing.

Mass spectrum of Intein-PPSIG
% Intensity
Mass (m/z)
4.0E+4
100
90
80
70
60
50
40
30
20
10
0
499
1402
2305
3208
4111
5014
566.22
684.20
758.36
855.43
1041.52
1174.61
1285.59
1301.58
1389.66
1447.63
1474.64
1534.74
1598.67
1748.83
1754.79
1766.91
1892.91
2024.07
2123.03
2146.00
2280.06
2302.05
2377.11
2475.08
2550.24
2632.23
2758.21
2886.30
2965.50
3339.66
3449.54
3551.08
3876.97
4324.69
4563.09

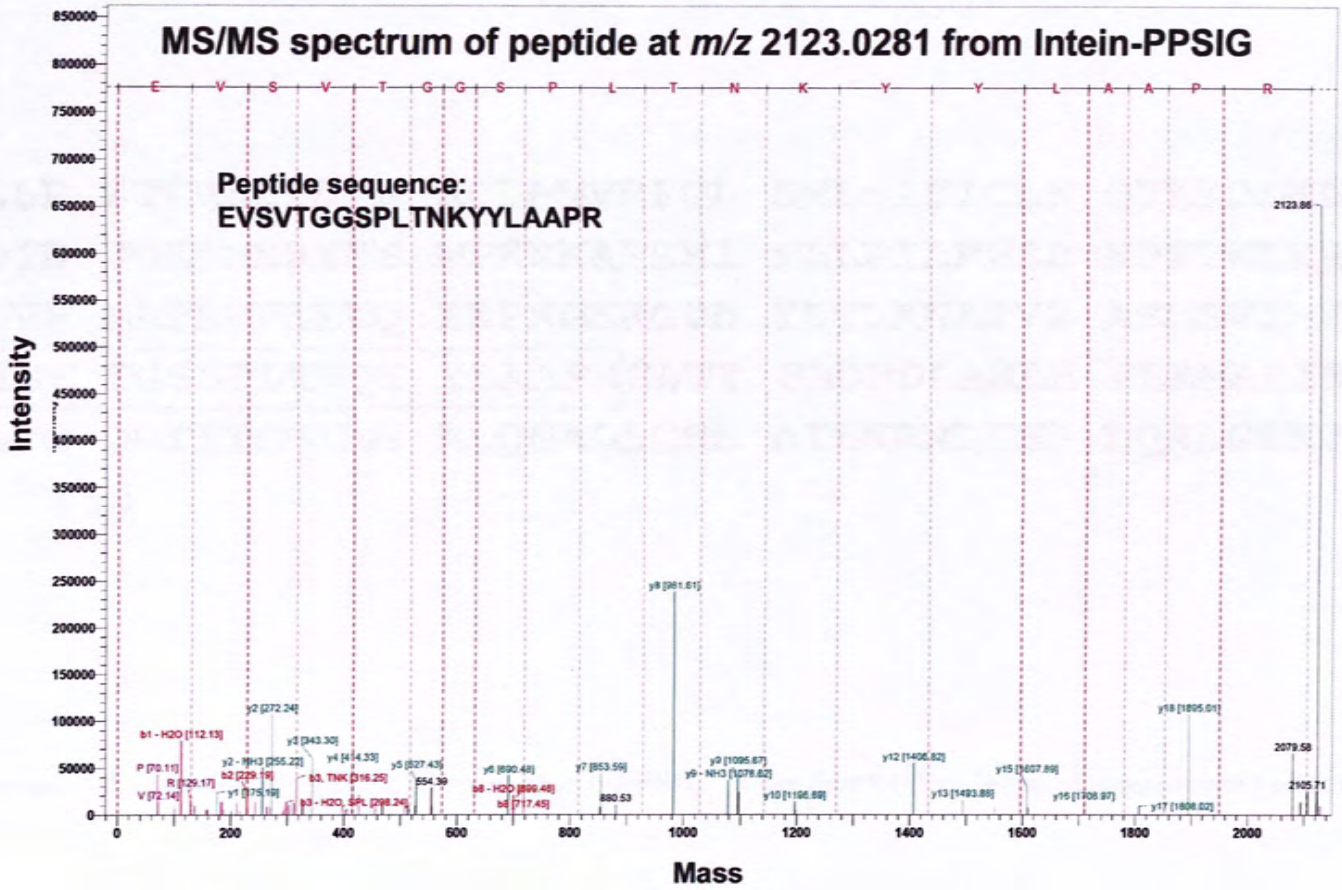


Figure 3.15 MS/MS spectrum of peptide peak at *m/z* 2123.0 from Intein-PPSIG

MS/MS spectrum generated for the peptide peak at *m/z* 2123.0 gave a sequence of EVSVTGGSPLTNKYYLAAPR, which matched with the amino acid sequence of PPSIG.

MFNTYQHLLP ETVRHLPDVK KQLAMVRPGL SMLSIFICLK GTKEDLKLQS 50
TNYYVYFDTD VDKAMERYVS MPKEKAPEHI PLLFIAFPSS KDPTWEERFP 100
DRSTMTALVP MAFEWFEEWQ EEPKGKRGVD YETLKNAFVE ASMSVIMELF 150
PRLEGKVESV TGGSPLTNQY YLAAPRGATY GADHDLARLH PHAMASIRAQ 200
TPIPNLYLTG QDIFTCGLMG ALQGALLCSS AILKRNLYSD LQALGSKVKA 250
QKKKM 255

Figure 3.16 Position of the peptide at *m/z* 2123.0 from Intein-PPSIG on the deduced amino acid sequence of the partial PPSIG for cloning

The peptide at *m/z* 2123.0 was located at the 157-176 region of the partial PPSIG used in cloning of Intien-PPSIG as shown in the sequence.

3.4 Preparation and characterization of Thio-PPSIG and Intein-PPSIG antisera

To minimize bacterial contaminants from PPSIG fusion protein antigens, the fusion proteins were resolved on preparative SDS-PAGE. Coomassie blue staining showed a major band at ~43 kDa for Thio-PPSIG (Figure 3.17) and ~87 kDa for Intein-PPSIG (Figure 3.18), and the protein bands were excised and homogenized in order to recover the antigen. After checking of antigen purity with mini SDS-PAGE, five boostings were made for each antigen. To determine the antigenicity of the antisera against PPSIG, the 5^{th} bleeding antisera and pre-immune serum were used as primary antibodies in Western blottings of recombinant Thio-PPSIG and Intein-PPSIG. Western blottings of Thio-PPSIG protein with Thio-PPSIG antiserum gave a strong immunoreactive band at ~43 kDa, which was absent with pre-immune antiserum (Figure 3.19), suggesting the immunoreactivity was due to immunization with Thio-PPSIG. The thioredoxin tag in the pThioHis vector control lane was also lighted up, which was in agreement with the fact that thioredoxin protein was present on the antigen as a fusion partner, which lead to the generation of non-specific antibodies. As Intein-PPSIG antiserum is not antigenic to thioredoxin, only Thio-PPSIG but not thioredoxin was lighted up as a ~43 kDa protein using the same protein samples in the Western blotting. Similarly, antiserum against Intein-PPSIG

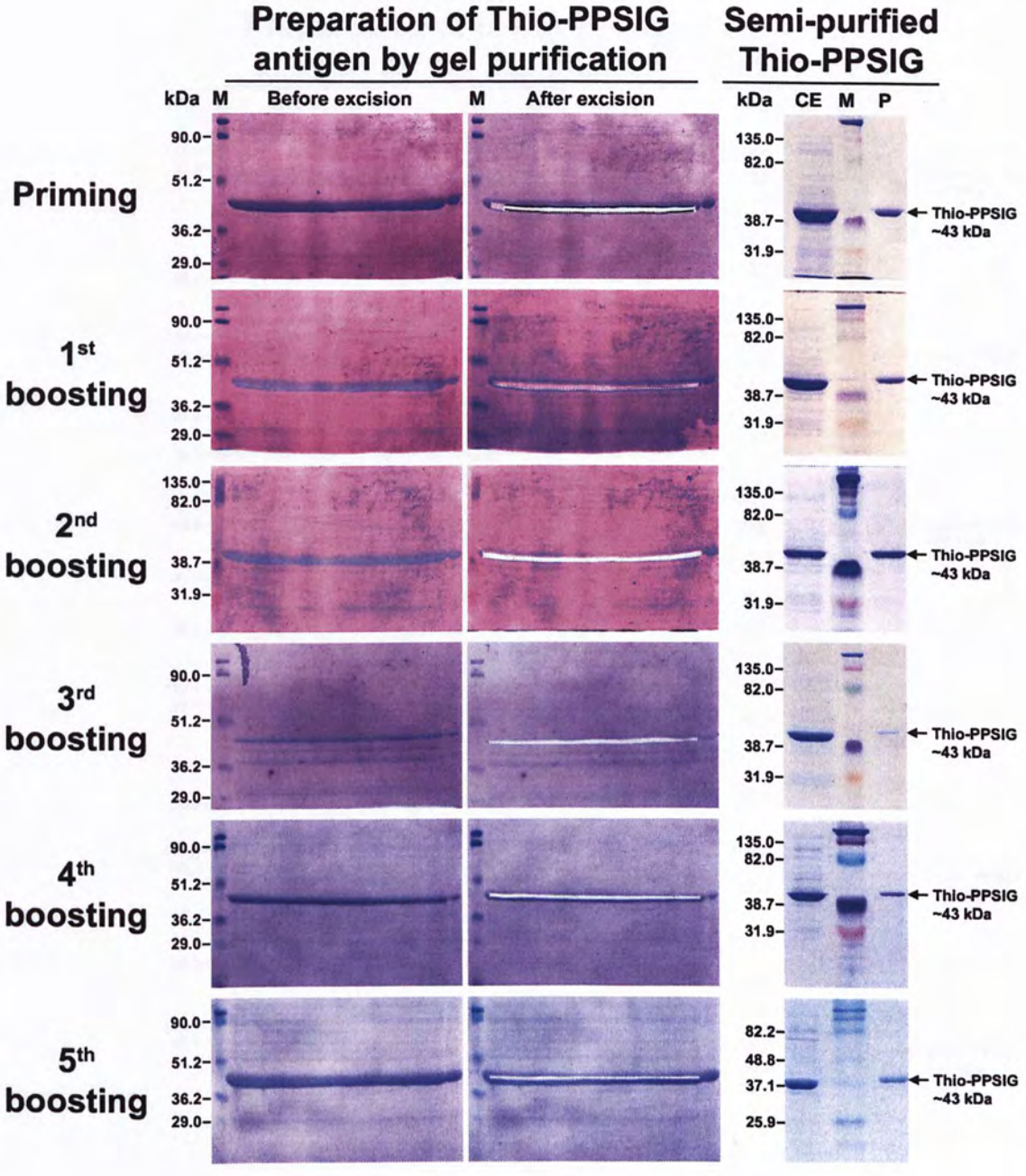


Figure 3.17 Preparation of Thio-PPSIG antigen for rabbit immunization (priming to fifth boosting)

Electrophoresis of ~1 mg of column-eluted refolded recombinant Thio-PPSIG on a 10% SDS-PAGE gel for gel excision and homogenization in PBS. Semi-purified Thio-PPSIG, after checking for purity, was injected into rabbits for different boostings. CE, column-eluted refolded Thio-PPSIG. P, semi-purified Thio-PPSIG. M, SDS-PAGE protein standards.

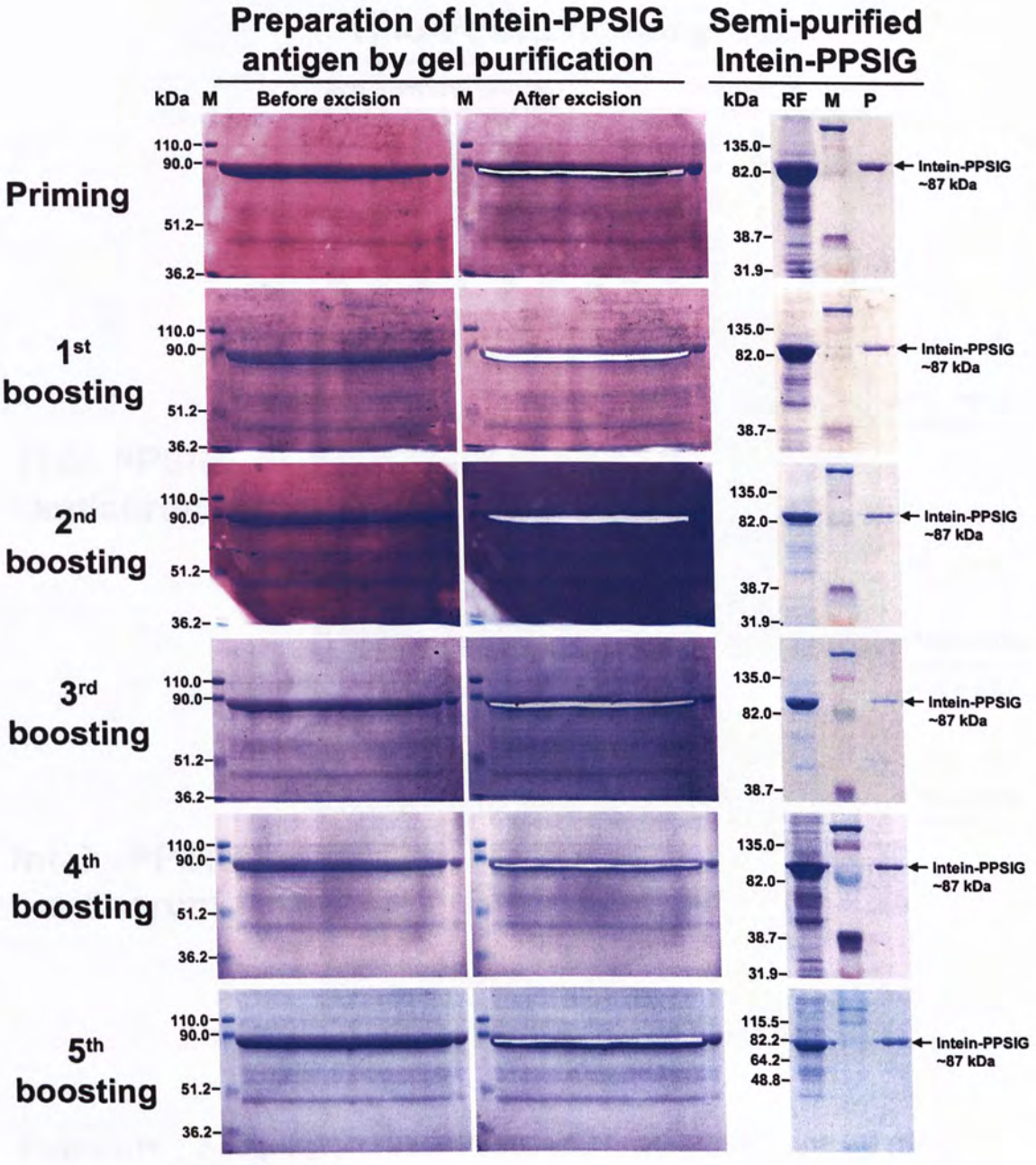


Figure 3.18 Preparation of Intein-PPSIG antigen for rabbit immunization (priming to fifth boosting)

Electrophoresis of ~1 mg of refolded recombinant Thio-PPSIG on an 8% SDS-PAGE gel for gel excision and homogenization. Semi-purified Thio-PPSIG, after checking for purity, was injected into rabbits for different boostings. RF, refolded Intein-PPSIG. P, semi-purified Intein-PPSIG. M, SDS-PAGE protein standards.

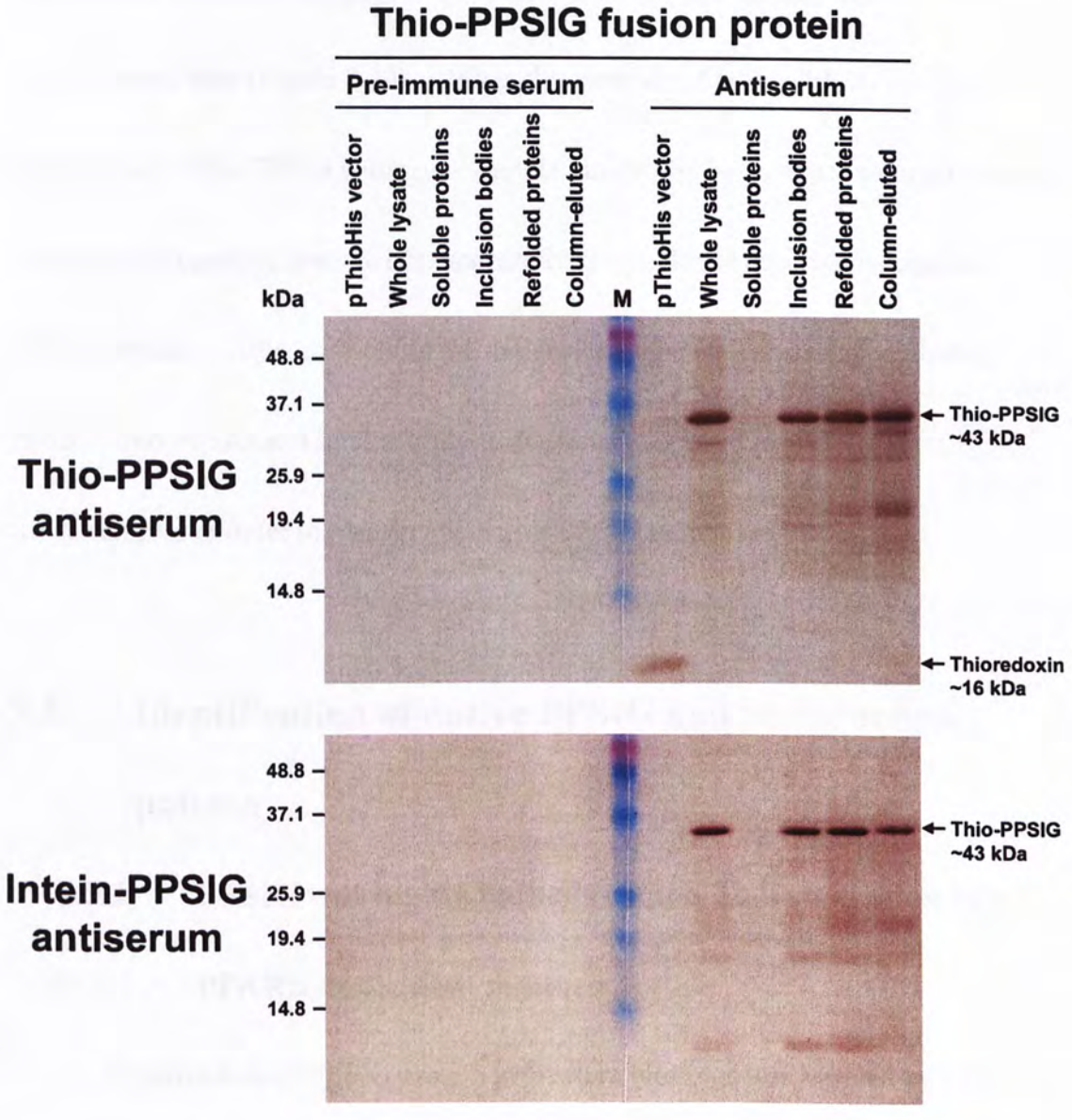


Figure 3.19 Antigenicity of Thio-PPSIG and Intein-PPSIG antisera (fifth bleeding) against Thio-PPSIG antigen

Western blotting of different fractions of bacterial lysate containing Thio-PPSIG detected with 1:5000 Thio-PPSIG and Intein-PPSIG antisera. Thio-PPSIG lighted up an immunoreactive band of ~43 kDa corresponding to Thio-PPSIG in whole lysate, inclusion bodies, refolded proteins and column-eluted proteins, as well as the ~16 kDa thioredoxin in the control lane transformed with pThioHis vector, which were absent with pre-immune serum. Intein-PPSIG antiserum lighted up the same immunoreactive band but not thioredoxin. M, prestained protein standards.

successfully detected both Intein-PPSIG fusion protein and intein protein in the pTYB vector control lane (Figure 3.20). When the same set of Intein-PPSIG protein was detected with Thio-PPSIG antiserum, only Intein-PPSIG but not the intein protein was lighted up, suggesting that the immunoreactivity was due to antigenicity against PPSIG protein. After confirming the antigenicity against bacterial recombinant PPSIG, Thio-PPSIG and Intein-PPSIG antisera were applied in Western blottings of tissue samples in order to identify the native PPSIG in mouse liver.

3.5 Identification of native PPSIG and its induction pattern

3.5.1 PPSIG was highly inducible upon 72-h starvation in a PPARα dependent manner

Identification of PPSIG protein in Western blottings was assisted by observing its induction pattern by starvation in PPARα wild-type and knockout mice, which is expected to match with its mRNA induction pattern. Mouse liver, having a high constitutive expression level as well as a high induction level of PPSIG, was chosen for the investigation of both mRNA and protein induction. As a confirmation to previous findings in our lab, Northern blotting analysis of total liver RNA showed that PPSIG mRNA level was significantly induced upon 72-h starvation in wild-type mice

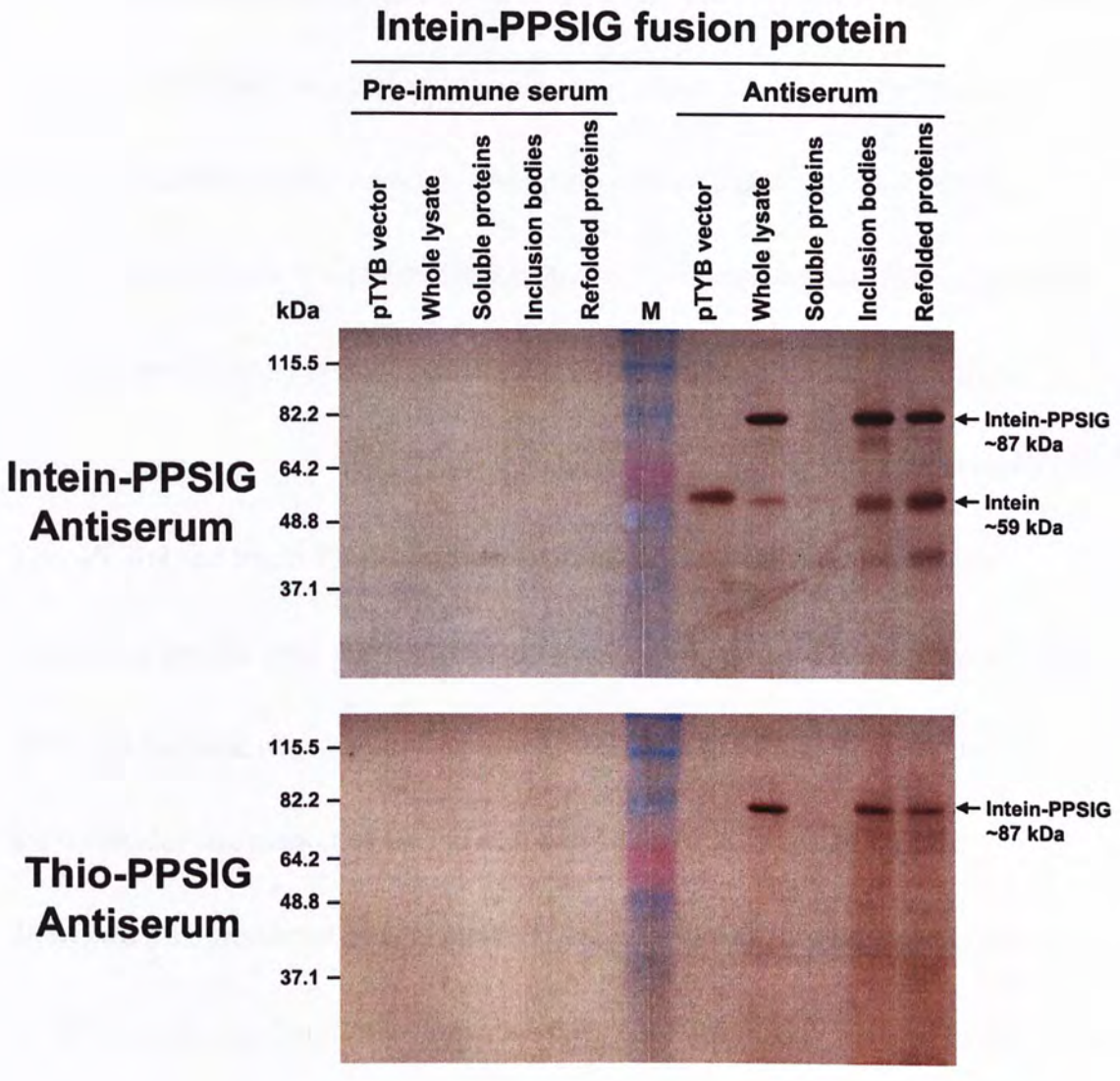


Figure 3.20 Antigenicity of Thio-PPSIG and Intein-PPSIG antisera (fifth bleeding) against Intein-PPSIG antigen

Western blottings of different fractions of bacterial-expressed Intein-PPSIG using Thio-PPSIG and Intein-PPSIG antisera with a titer of 1:5000. For Intein-PPSIG antiserum, detection of a ~87 kDa immunoreactive band corresponding to Intein-PPSIG was observed in whole lysate, incusion bodies and refolded proteins, accompanied by a ~59 kDa intein protein. For Thio-PPSIG antiserum, only the ~87 kDa immunoreactive band was lighted up. In each case, pre-immune serum was unable to light up any non-specific protein bands. M, prestained protein standards.

but not in knockout mice (Figure 3.21). The result was normalized with 18S and 28S ribosomal RNA, which showed similar intensity in all lanes. Moving on to the induction pattern of PPSIG protein, liver samples were subjected to subcellular fractionation to yield endoplasmic reticulum in microsomal fraction, which contained PPSIG as predicted by previous experiments (Sun, paper submitted). A strong immunoreactive band of ~60 kDa comparable to that of PPSIG was lighted up by both Thio-PPSIG and Intein-PPSIG antisera and was absent using pre-immune sera, suggesting that the band was contributed by antibodies against PPSIG (Figure 3.22). Although the band was smaller than the expected size of 67 kDa when compared with the molecular size marker of 64.2 kDa, it was believed to be due to the size discrepancy of prestained protein markers. Comparing the protein expression level in PPARα wild-type mice, the immunoreactive band was highly induced by starvation, consistently in all the three mice within a group. In the absence of PPARα, its constitutive expression level was lower than in wild-type mice and was slightly depressed by starvation in knockout mice. As the protein induction pattern was in coincidence with mRNA level in Northern blotting, the immunoreactive band was likely to represent native PPSIG. On the other hand, as the intensity of the immunoreactive band could be affected by the amount of microsomes loaded on each lane, it was monitored by calnexin, an integral membrane protein that localizes evenly

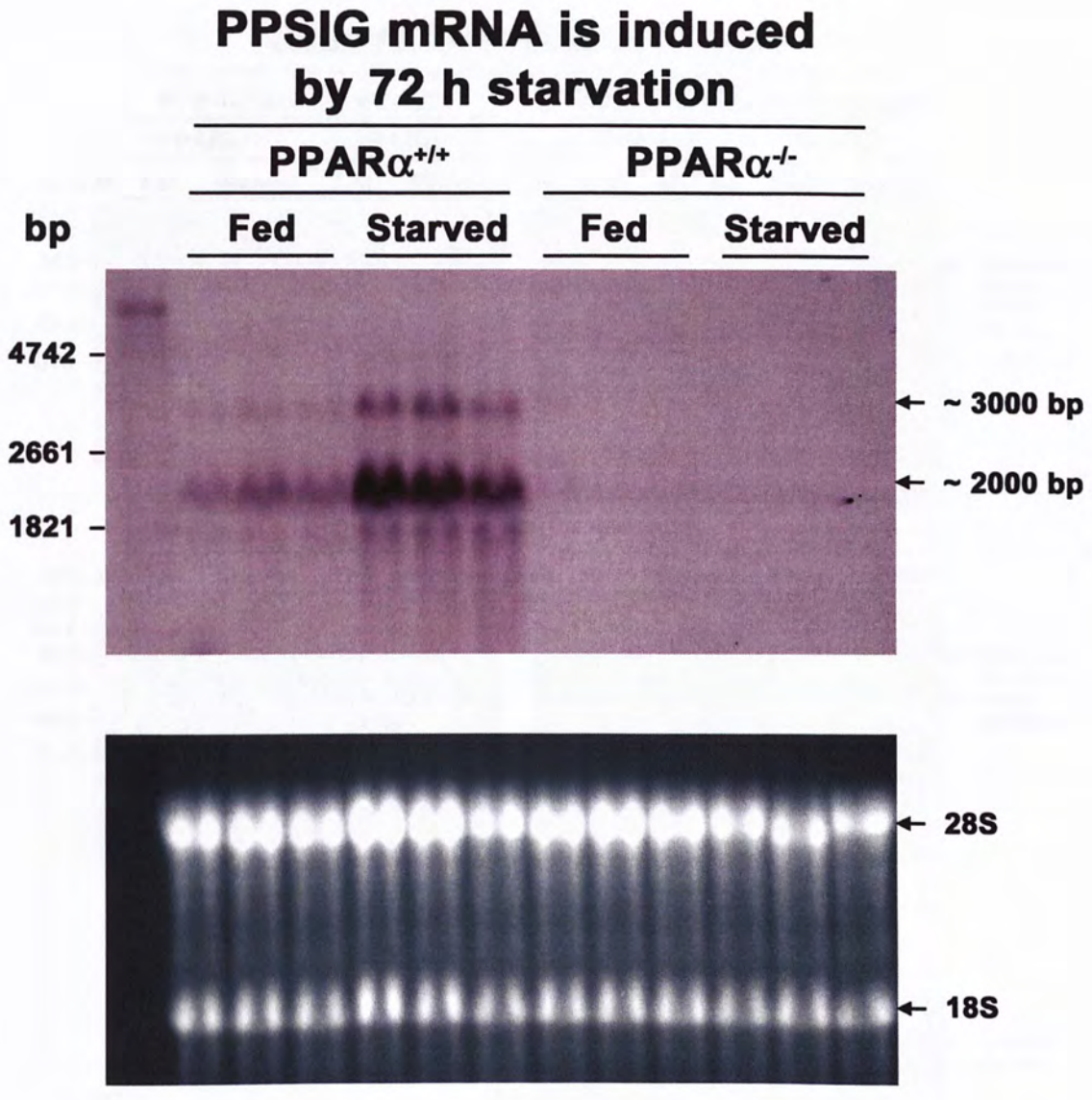


Figure 3.21 PPSIG mRNA induction in liver (starvation experiment) using Northern blotting analysis

Each lane contained 30 μg total liver RNA from PPARα wild-type and knockout mice either fed or starved for 72 h. The ~2000 bp band represented PPSIG mRNA, which had a stronger signal than its pre mRNA with a size of ~3000 bp. Level of both transcripts was significantly induced upon 72 h starvation in wild-type mice but not in knockout mice. The 18S and 28S ribosomal RNA was shown for normalization to ensure equal RNA loading in each well.

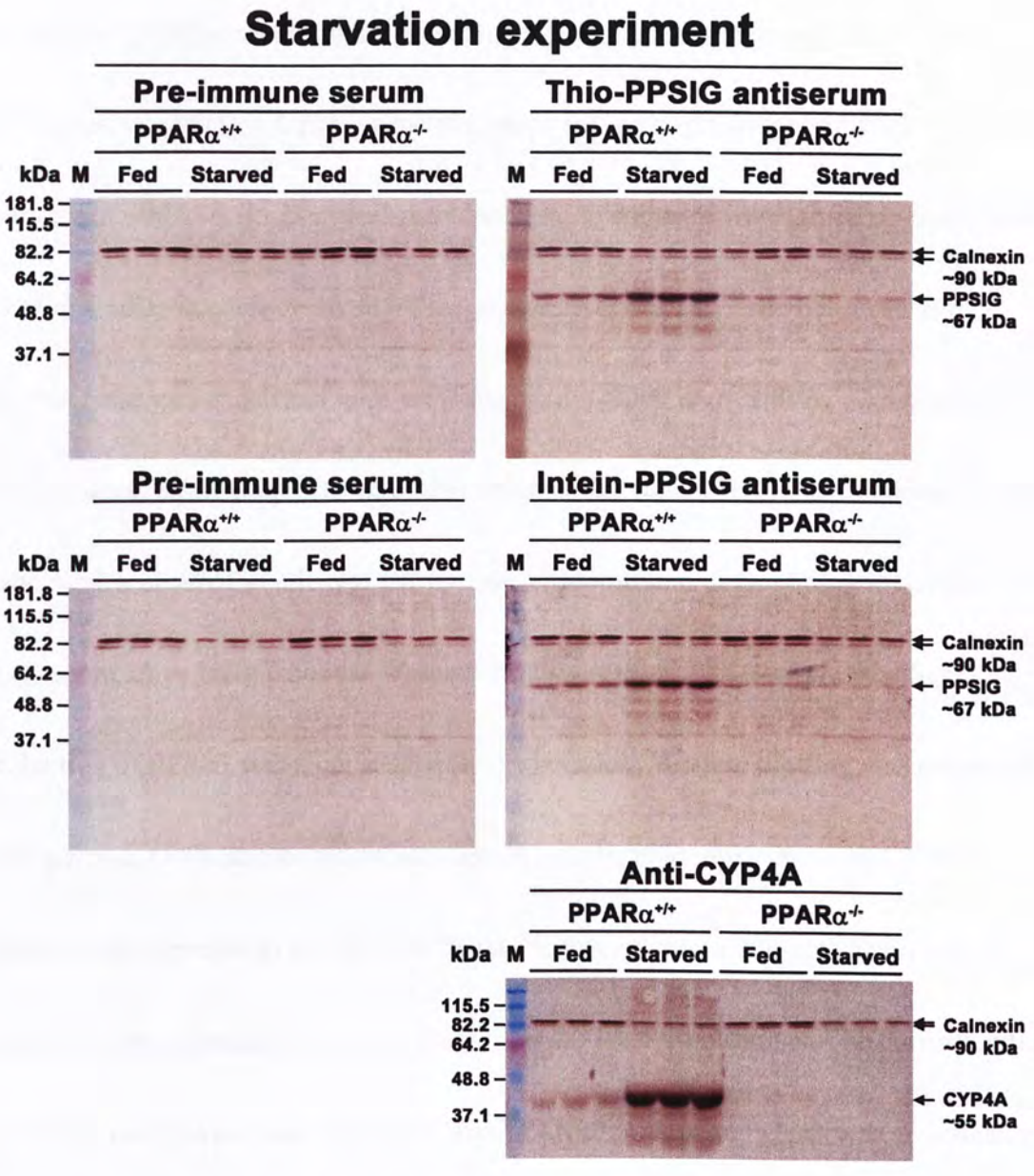


Figure 3.22 PPSIG protein induction in liver microsomes (starvation experiment) using Western blotting analysis

Each lane contained 40 μg of liver microsomes from PPARα wild-type and knockout mice either fed or starved for 72 h. Anti-calnexin antibody (1:2000) was co-incubated with pre-immune serum or antiserum, which lighted up a double band at ~90 kDa. A ~60 kDa immunoreactive band (expected to be ~67 kDa) was lighted up by both Thio-PPSIG and Interin-PPSIG antisera (1:1000) but not by pre-immune serum (1:500). The band was highly induced in starved wild-type mice. Similar pattern was observed using anti-CYP4A antiserum (1:2000), which lighted up the ~55 kDa CYP4A protein that is induced during starvation. M, prestained protein standards.

in the endoplasmic reticulum as performed by other research groups (Klett *et al.* 2004) (Waugh *et al.* 2003). Upon co-incubation of anti-calnexin antiserum with PPSIG antisera, a ~90 kDa double band corresponding to calnexin was lighted up in all lanes having similar intensity with slight inconsistent variations. The observation of a double band was consistent with previous results (Klett *et al.* 2004). Amount of microsomes, being proportional to that of calnexin, was therefore comparable in each lane, which ensured that it was fair to observe the protein expression level of the immunoreactive band from the Western blotting results. To observe whether induction of PPSIG was a consequence of starvation, Western blotting was performed using anti-CYP4A antiserum, which lighted up a band at ~40 kDa. As CYP4A protein was expected to be ~55 kDa in size, it further shows that prestained protein marker underestimated the size of protein on the Western blottings. The corresponding CYP4A protein was a well-characterized PPARα target gene, which was significantly upregulated upon starvation. As CYP4A was found to be induced in wild-type mice but not in knockout mice, it was confident to deduce that the condition of starvation is optimum in activating PPARα. It added to the deduction that induction of PPSIG was PPARα-dependent and causally related to starvation. To further investigate the PPARα-dependence of PPSIG induction, its expression pattern was observed in mice treated with Wy-14,643, an exogenous PPARα ligand.

3.5.2 PPSIG showed slight induction upon 2-wk Wy-14,643 treatment

Being a well-documented exogenous ligand that could activate PPARα more specifically than starvation, 2-wk treatment with Wy-14,643 triggered PPSIG induction in a different way. Similar to the induction pattern in starvation experiment, the mRNA level of PPSIG was highly increased upon 2-wk Wy-14,643 treatment in wild-type mice but not in knockout mice (Figure 3.23). However, PPSIG protein induction deviated from the induction pattern of mRNA. Consistent with both Thio-PPSIG antiserum and Intein-PPSIG antiserum, the same immunoreactive band was lighted up with similar intensity among the three mice in the same group (Figure 3.24). Similar to starvation treatment, constitutive expression level of PPSIG was higher in PPARα wild-type mice. However, the induction level of PPSIG was hardly observable in wild-type mice although its depression in knockout mice was consistent with starvation treatment. The credibility of PPSIG expression level was supported by the consistent level of calnexin double band in each lane. On the other hand, high induction level of PPSIG mRNA as well as the CYP4A protein in wild-type mice eliminated the possibility that Wy-14,643 treatment did not activate PPARα and thus could not induce PPSIG.

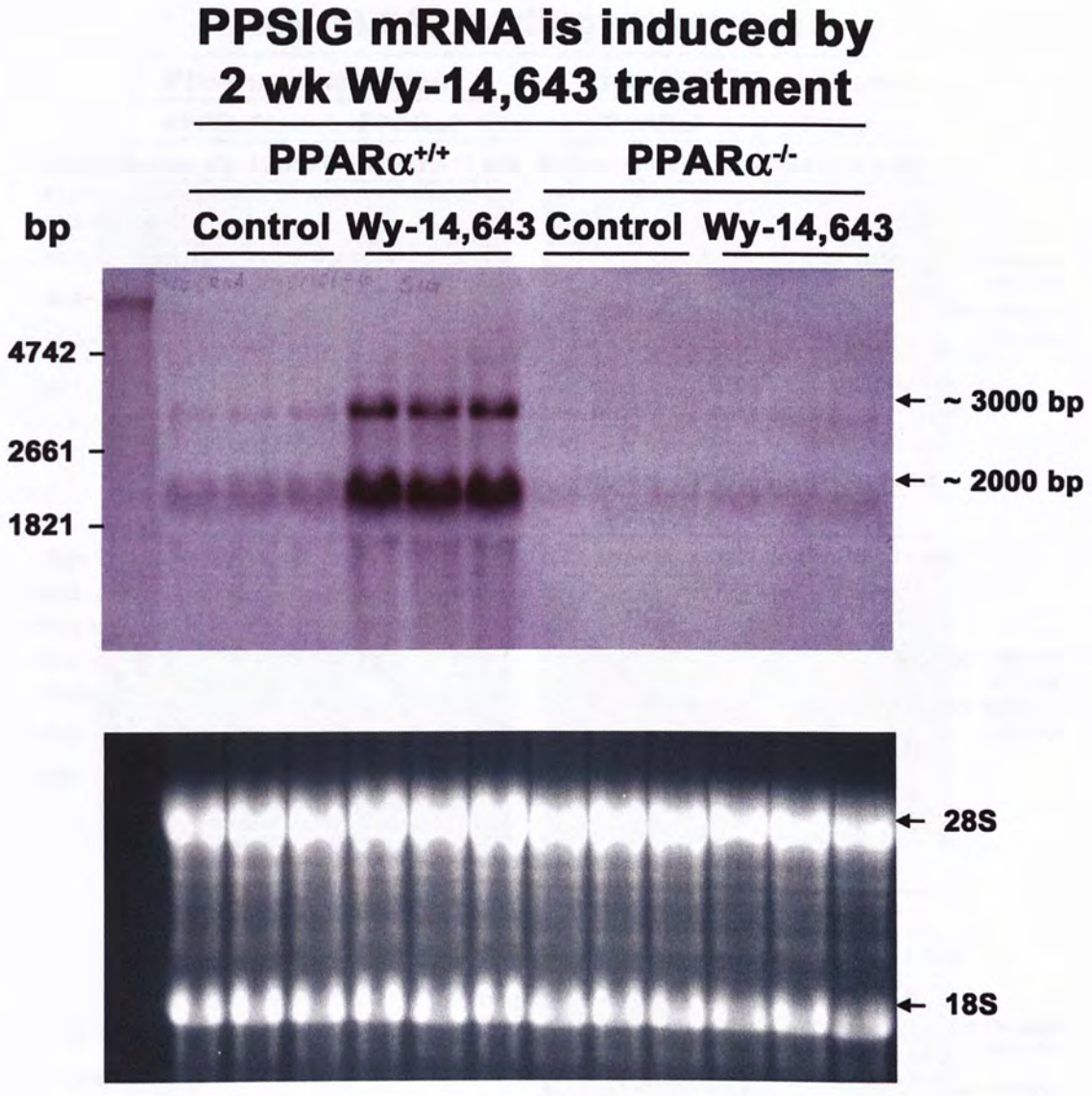


Figure 3.23 PPSIG mRNA induction in liver (Wy-14,643 feeding experiment) using Northern blotting analysis

Each lane contained 30 μg total liver RNA from PPARα wild-type and knockout mice either fed with control diet or 0.1% (w/w) Wy-14,643 for 2 wk. Both ~2000 bp and ~3000 bp bands were highly induced by 2 wk Wy-14,643 treatment in wild-type mice. The same induction was not observed for knockout mice. The 18S and 28S ribosomal RNA showed that similar amount of RNA was loaded in each well.

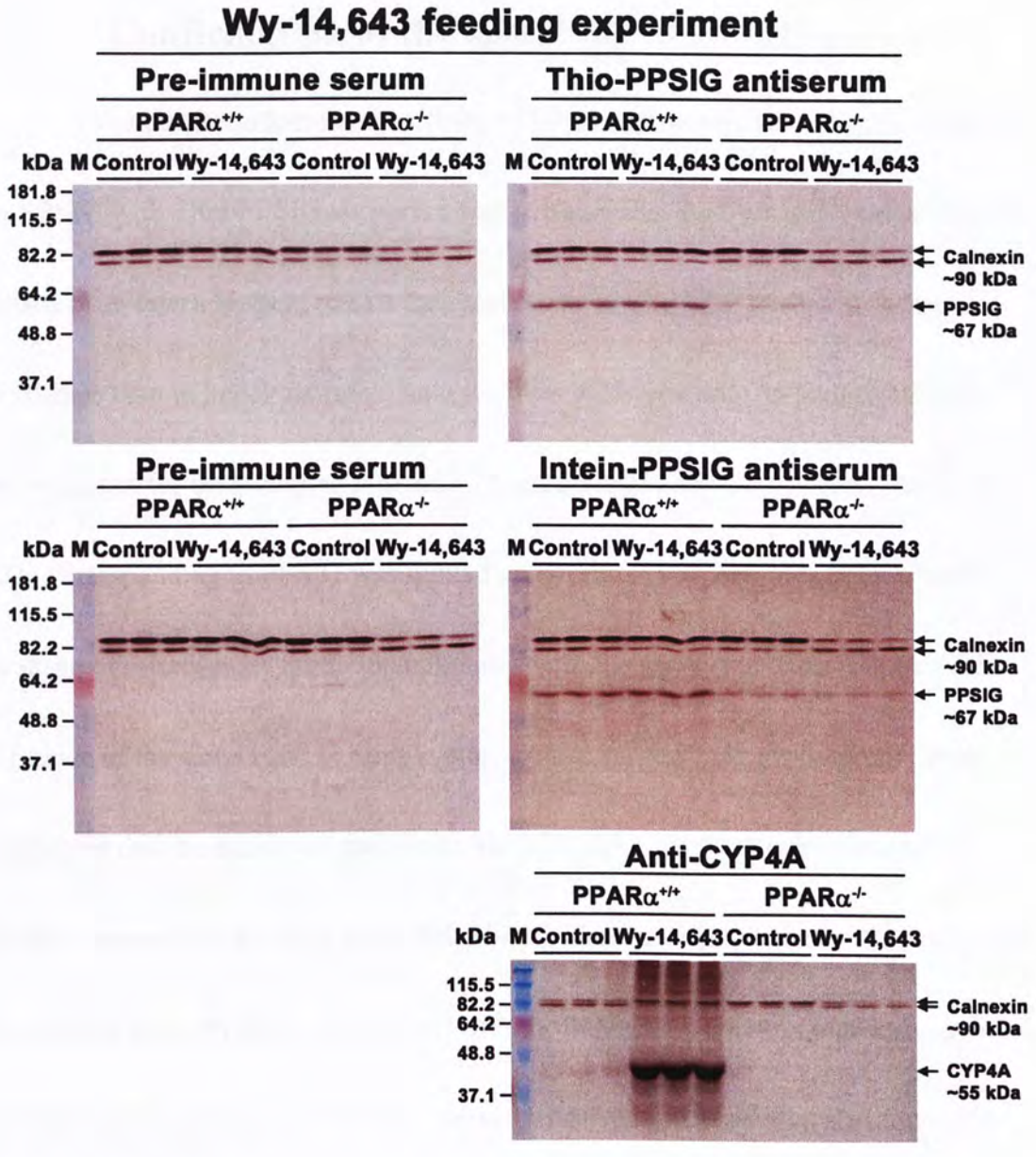


Figure 3.24 PPSIG protein induction in liver microsomes (Wy-14,643 feeding experiment) using Western blotting analysis

Each lane contained 40 μg of liver microsomes from PPARα wild-type and knockout mice either fed with control diet or 0.1% (w/w) Wy-14,643 for 2 wk. Anti-calnexin antibody (1:2000) lighted up a double band at ~90 kDa, when co-incubated with pre-immune serum or antiserum. A ~60 kDa immunoreactive band (expected to be ~67 kDa) was lighted up by both Thio-PPSIG and Interin-PPSIG antisera (1:1000) but not by pre-immune serum (1:500). No significant induction was observed in Wy-14,643 treated wild-type mice, despite high induction of the ~55 kDa CYP4A protein lighted up by anti-CYP4A antiserum (1:2000). M, prestained protein standards.

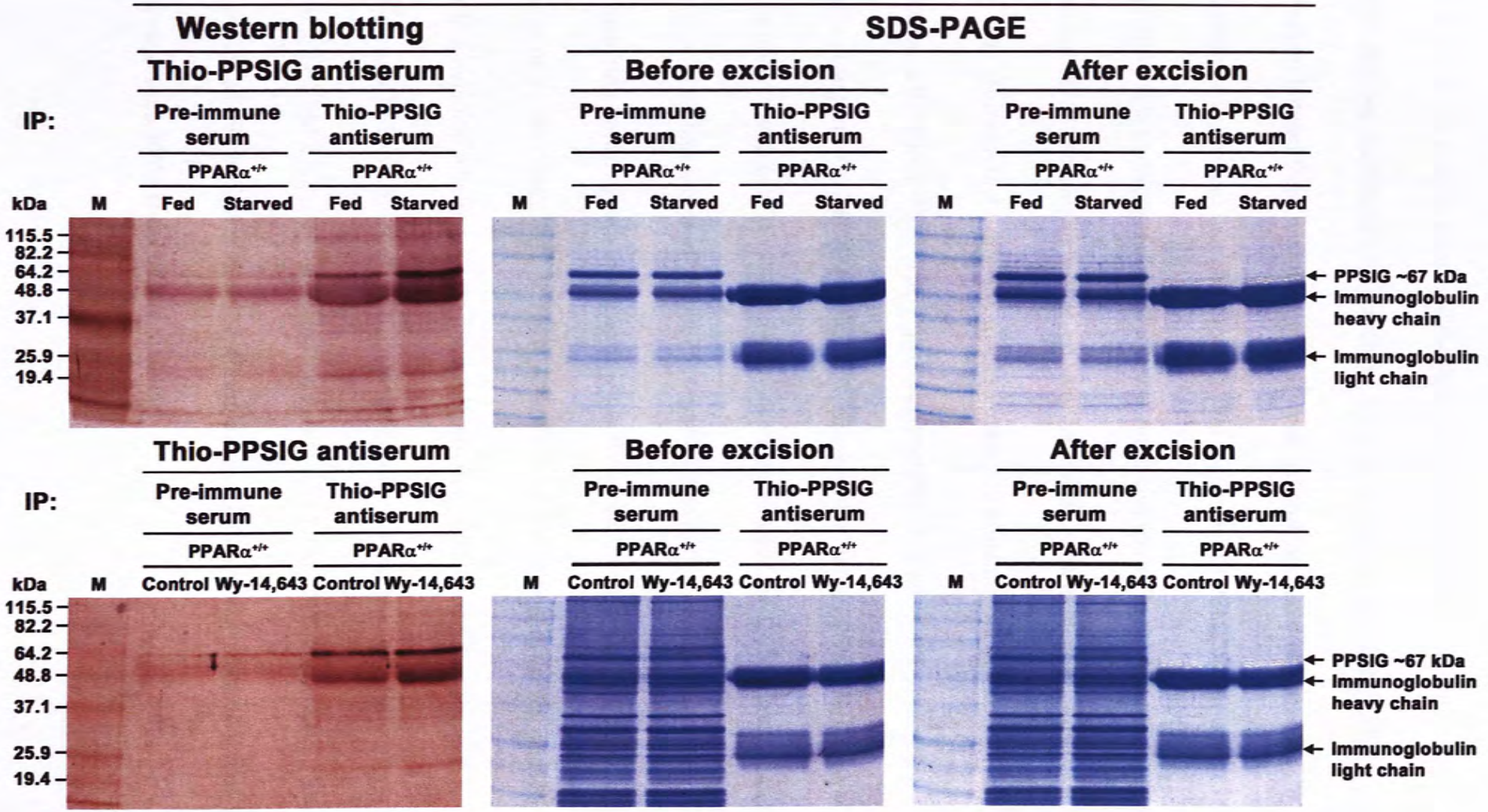
3.6 Confirmation of the specificity of PPSIG antiserum

In order to confirm the specificity of PPSIG antiserum, immunoprecipitation of PPSIG with Thio-PPSIG antiserum was performed using liver microsomal fraction. Based on Western blotting results that expression level of PPSIG was higher in wild-type than in knockout mice, samples from wild-type mice in both treatments were chosen for immunoprecipitation. A strong immunoreactive band of size ~67 kDa corresponding to PPSIG was lighted up by Thio-PPSIG antiserum in Western blotting of microsomes immunoprecipitated with the same antiserum (Figure 3.25). Absence of the same band in sample immunoprecipitated with pre-immune serum suggested that the band was specific to Thio-PPSIG. Two proteins bands below PPSIG represented the immunoglobulin heavy chain and light chain of the antibodies that pulled down PPSIG. Consistent with the induction pattern in total microsomes, PPSIG was highly induced by starvation in wild-type mice and slightly induced by Wy-14,643 treatment. In order to confirm the identity of the immunoreactive band being PPSIG, mass spectrometry was performed. With reference to the position of PPSIG on the Western blotting membrane, the protein band slightly above the immunoglobulin heavy chain was excised from the gel. After in-gel trypsin digestion and peptide extraction, the peptides recovered were analyzed by MALDI-TOF mass spectrometry, and mass spectra of the peptides were obtained. In

Figure 3.25 Immunoprecipitation of native PPSIG from liver microsomes

Native PPSIG was immunoprecipitation from liver microsomes of wild-type mice in starvation and Wy-14,643 exposure experiments. A 500 µg of microsomes from each group of mice was immunoprecipitated first with pre-immune serum (1:500) and then Thio-PPSIG antiserum (1:100). The immunoprecipitation products were resolved on an 8-18% gradient SDS-PAGE and transferred to a PVDF membrane. Western blottings with Thio-PPSIG antiserum lighted up a strong immunogenic band at ~67 kDa just above the immunoglobulin heavy chain. The protein band was located on a Coomassie blue stained SDS-PAGE gel loaded with the same set of sample with the help of Western blotting result. The band was excised into gel discs for msss spectrometry. M, prestained protein standards.

Immunoprecipitated PPSIG from liver microsomes



order to match the peptide masses in the mass spectra with that of a known protein, the PMF data were submitted for searching in the NCBI database. For PPSIG immunoprecipitated from the microsome of mice fed with normal diet, the search result returned no protein that was similar to the sequence of PPSIG, which was explained by low abundance of PPSIG in the sample, as shown in the Western blotting. Amount of PPSIG was higher in starved samples and the protein identification gave a match with all-*trans*-13,14-dihydroretinol saturase at a protein score of 74 (significance threshold=78). The 9 matching peptides covered 18.1% of the total amino acid sequence of PPSIG (Table 3.3). On the other hand, PPSIG from control mice or mice treated with Wy-14,643 were identified as all-*trans*-13,14-dihydroretinol saturase with protein scores of 78 and 75 respectively (significance threshold=78). Nine matching peptide peaks were identified in the control sample showing sequence coverage of 19.0% (Table 3.4). The Wy-14,643 treated sample also gave 9 peptide peaks that matched with the trypsin digestion pattern of PPSIG, and the sequence coverage was 17.1% (Table 3.5). The identification of the immunoprecipitated protein to be PPSIG not only supported the specificity of Thio-PPSIG antiserum, but also confirmed that PPSIG is an authentic protein in mice liver that could be induced by starvation in a PPARα-dependent manner.

Table 3.3 Nine peptide peaks matched with all-*trans*-13,14-dihydroretinol saturase using MALDI-TOF peptide mass fingerprinting after trypsin digestion of native PPSIG in liver microsomes from PPARα wild-type mice starved for 72 hours (starvation experiment)

Calculated mass	Observed mass	Position	Amino acid sequence
534.267	534.2275	453-456	FPDR
706.4246	706.3586	224-229	FGLLTR
1151.642	1151.5828	38-47	RPPEPLVTDK
1246.5811	1246.5358	531-542	GATYGADHDLAR
1308.6794	1308.6113	590-601	NLYSDLQALGSK
1316.7168	1316.66	317-329	ATVQSVLLDSAGR
1725.9282	1725.8768	293-308	GGSSEIAFHTIPLIQR
1758.9344	1758.8909	236-252	ASTQSLAEVLQQLGASR
2123.0767	2123.0298	511-530	VESVTGGSPLTNQYYLAAPR

Table 3.4 Nine peptide peaks matched with all-*trans*-13,14-dihydroretinol saturase using MALDI-TOF peptide mass fingerprinting after trypsin digestion of native PPSIG in liver microsomes from PPARα wild-type mice fed with control diet (Wy-14,643 exposure experiment)

Calculated mass	Observed mass	Position	Amino acid sequence
706.4246	706.3680	224-229	FGLLTR
1132.6045	1132.5497	543-552	LHPHAMASIR
1151.642	1151.5988	38-47	RPPEPLVTDK
1246.5811	1246.5647	531-542	GATYGADHDLAR
1308.6794	1308.6378	590-601	NLYSDLQALGSK
1316.7168	1316.6489	317-329	ATVQSVLLDSAGR
1725.9282	1725.9010	293-308	GGSSEIAFHTIPLIQR
1758.9344	1758.8802	236-252	ASTQSLAEVLQQLGASR
2123.0767	2123.0203	511-530	VESVTGGSPLTNQYYLAAPR

Table 3.5 Nine peptide peaks matched with all-*trans*-13,14-dihydroretinol saturase using MALDI-TOF peptide mass fingerprinting after trypsin digestion of native PPSIG in liver microsomes from PPARα wild-type mice fed with 0.1% (w/w) Wy-14,643 for 2 weeks (Wy-14,643 exposure experiment)

Calculated mass	Observed mass	Position	Amino acid sequence
534.267	534.2142	453-456	FPDR
706.4246	706.3575	224-229	FGLLTR
932.4108	932.3807	446-452	DPTWEER
1132.6045	1132.5415	543-552	LHPHAMASIR
1246.5811	1246.5408	531-542	GATYGADHDLAR
1308.6794	1308.623	590-601	NLYSDLQALGSK
1725.9282	1725.8755	293-308	GGSSEIAFHTIPLIQR
1758.9344	1758.8807	236-252	ASTQSLAEVLQQLGASR
2123.0767	2123.0203	511-530	VESVTGGSPLTNQYYLAAPR

Chapter 4 Discussion

PPSIG is a novel PPARα target gene discovered in our laboratory, and current characterization of its functional role is directed by its PPARα-dependent inducibility. As its mRNA level was highly induced by either 72-h starvation or 2-wk Wy-14,643 treatment as shown in FDD and Northern blottings, it is important to examine its induction pattern at protein level to determine if the protein levels match with the mRNA expression patterns in the starvation and Wy-14,643 treatments. By expressing recombinant PPSIG fusion proteins for rabbit immunization, we have raised polyclonal antisera against PPSIG, which can be used to identify the PPSIG protein and study its induction pattern in mouse liver tissue.

To increase the chance of getting a specific antiserum, the immunization strategy involved two different antigens. Thioredoxin tag was added at the N-terminal of PPSIG while intein tag was added at the C-terminal. Thioredoxin was comparable in size to the truncated PPSIG while intein was much larger than PPSIG. Each fusion partner was expected to induce non-specific antibodies in the antiserum, as confirmed with the Western blotting results that the antisera lighted up also the lane transformed with vector only. The antigenicity against PPSIG was observed by the

immunoreactivity of each antiserum towards the recombinant protein with a different fusion partner. Thio-PPSIG antiserum lighted up Intein-PPSIG antigen and *vice versa*, suggesting that a considerable amount of antibodies against PPSIG was present in the antisera. It further showed that the size and position of the fusion partner did not obviously affect the immune response against PPSIG. Nevertheless, non-specific antibodies against the fusion partner limited the use of the antisera on experiments requiring a highly specific antibody, such as immunohistochemistry and affinity purification of PPSIG protein. If these experiments are to be carried out in the future, it may be better to obtain an antiserum using pure PPSIG protein as antigen. However, non-specific antibodies do not greatly affect the result in the current project. As the antisera are shown to be antigenic to PPSIG, it is suitable to apply them on the identification of native PPSIG in mouse liver.

The native PPSIG was identified as a ~67 kDa protein in liver microsomes, where PPSIG is predicted to localize in. Its exact size would further be confirmed by N-terminal sequencing, in order to determine whether the signal peptide is retained at its N-terminal. Despite the presence of a fusion partner on its antigen, PPSIG was lighted up as a strong major band, indicating the specificity of the antisera. Detection of occasional minor bands on the Western blottings was either contributed

by the non-specific antibodies against the fusion partner or proteins having a similar conformation to PPSIG. Observation of PPSIG was assisted by the induction pattern by starvation and Wy-14,643 in PPARα wild-type and knockout mice. As the mRNA expression can be induced by both starvation and Wy-14,643 feeding, it was expected that the protein induction pattern would match with mRNA level. PPSIG protein was obviously induced by starvation in PPARα, which strengthens the conclusion that the ~67 kDa immunogenic band was PPSIG. However, the protein induction pattern by Wy-14,643 deviated from the mRNA induction pattern although the same protein was identified from the Western blotting. Only slight induction of PPSIG was observed in wild-type mice fed with Wy-14,643 for 2 wk.

To confirm the immunogenic band as PPSIG, mass spectrometry was performed followed by immunprecipitation with Thio-PPSIG antiserum. Being well-characterized at its mRNA level, there has not been solid identification of native PPSIG based on it protein sequence. Mass spectrometry analysis followed by immunoprecipitation gave a hit of all-*trans*-13,14-dihydroretinol saturase in 3 out of 4 samples. PPSIG was identified in wild-type mice starved for 72 h, fed with control diet and fed with 0.1% (w/w) Wy-14,643 for 2 wk. The major limitation of immunoprecipitation was the contamination of PPSIG with rabbit immunoglobulin

heavy chain, which is especially significant for PPSIG having similar size with the heavy chain. In addition, low amount of PPSIG recovered after immunoprecipitation makes MS/MS sequencing not feasible and resulted in a relatively low protein score. Nevertheless, the consistent identification of PPSIG in 3 samples mentioned above should lead to a reliable conclusion that the major immunogenic band lighted up by PPSIG antisera corresponded to PPSIG. It provides substantial proof that PPSIG, with its protein sequence derived from its mRNA, is not only a hypothetical protein, which further supports our interpretation of PPSIG induction in mouse liver microsomes.

PPSIG, being found to be inducible by starvation in a PPARα-dependent manner at its mRNA level, is believed to play an important role in lipid metabolism. Having its starvation inducibility confirmed by Western blotting, it is reasonable to predict its function with reference to a starved condition, where lipids are broken down into fatty acids for energy production. Although starvation also triggers a change in protein expression profile in the body, a direct relationship should be drawn between PPSIG induction and generation of fatty acids, which are endogenous PPARα ligands. A number of PPARα-regulated genes with different functions have been reported to be inducible by starvation and take part in lipid metabolism. Those

having a peroxisome proliferator response element report are usually inducible also by exogenous PPARα ligands such as fibrates, so the protein induction pattern of PPSIG by Wy-14,643 was investigated.

Unexpectedly, treatment of Wy-14,643 failed to induce expression of PPSIG protein. From Northern blotting results, the mRNA level of PPSIG was induced by 2-wk treatment with Wy-14,643, which belongs to a group of hypolipidemic drug used to lower blood cholesterol level. Being structurally similar to fatty acids, it is a ligand for PPARα that activates the transcription of PPARα target genes. However, failure of causing a corresponding induction of PPSIG protein implied that the translational and post-translational steps involve a more complicated mechanism. This phenomenon has been reported for carbonic anhydrase IX, which was suggested to have low translation efficiency (Hilvo *et al.* 2004). It is not likely for PPSIG, as there is considerable expression of PPSIG in PPARα wild-type mice fed with normal diet. However, it is possible that excess mRNA is degraded by microRNAs, which has been shown to regulate the expression of a number of genes (Carthew 2006). Induction of PPSIG protein during starvation but not Wy-14,643 treatment suggested that the induction of microRNAs against PPSIG mRNA may be dependant on the presence of Wy-14,643. However, as the biological role and mechanism of

microRNAs are not fully characterized, it would be difficult to draw a conclusion on whether microRNAs did take part in the regulation of PPSIG mRNA level. In order to further examine the effect of Wy-14,643 treatment to PPSIG protein induction, feeding the mice with different dosage of Wy-14,643 for different time periods may be considered. On the other hand, localization to other cellular compartments is not expected, as shown by GFP-PPSIG fusion protein in mammalian cells. As PPSIG was predicted to contain a secretory signal peptide, it is possible that the PPSIG synthesized did not stay in endoplasmic reticulum and is secreted outside the cells. It is also possible that the nascent PPSIG chain is misfolded and being degraded by the quality control machinery in the endoplasmic reticulum (Ellgaard and Helenius 2003). Degradation and removal of misfolded protein is assisted by the formation of inter-chain disulphide bonds (Molinari *et al.* 2002), which may also contributed to the formation of inclusion bodies during expression in *E. coli*. It is not surprising because Wy-14,643 was applied to specifically activate PPARα, and may cause the expression of different proteins that destabilizes or aggregate PPSIG. No matter why PPSIG protein failed to be induced by Wy-14,643 treatment, it is strongly believed that PPSIG was an important protein during starvation. In order to predict the role of PPSIG during starvation, it is important to compare its induction pattern and subcellular localization with other well-characterized PPARα target genes.

PPSIG shows similar induction pattern to pyruvate dehydrogenase kinase 4 (PDK-4). While its mRNA level can be induced by both starvation and Wy-14,643 treatment, its protein induction level is much higher during starvation (Holness *et al.* 2003; Sugden *et al.* 2002). It has been extensively studied to inhibit the activity of pyruvate dehydrogenase, a key enzyme in glucose metabolism which catalyzes the conversion of pyruvate to acetyl-CoA for production of energy in the TCA cycle. Therefore, activation of PPARα shifts the energy homeostasis equilibrium from glucose oxidation to gluconeogenesis, sparing glucose for utilization by the brain. Considering the biological function of PDK-4, starvation would be a more reasonable inducer, as in the case of PPSIG. However, PDK-4 differs from PPSIG in terms of its tissue distribution. It is highly inducible in skeletal muscle after 24 hr of starvation and decline gradually afterwards (Li *et al.* 2006). In addition, it can be induced by high fat diet (Holness *et al.* 2000), which does not affect the expression of PPSIG (data unpublished).

Another PPARα-inducible gene is the acyl-CoA oxidase, which catalyzes the dehydrogenation of fatty acyl-CoA to *trans*-2-enoyl-CoA, a committing step into peroxisomal β-oxidation. It helps the production of energy from fatty acids during starvation. With the identification of a peroxisome proliferator responsive element,

its induction in both mRNA and protein level is believed to be PPARα-mediated (Tugwood *et al.* 1992). Its hydrogen withdrawing ability may be due an FAD binding site (Tokuoka *et al.* 2006), which is also present in PPSIG as predicted from its amino acid sequence. Therefore, it is likely for PPSIG to possess dehydrogenase activity. However, as PPSIG is localized in the endoplasmic reticulum, it is necessary to compare its inducibility with an enzyme from the same subcellular compartment.

The cytochrome P450 4A enzymes localize extensively in the smooth endoplasmic reticulum and are responsible for fatty acid ω-hydroxylation (He *et al.* 2005). After oxidizing the carbon most distant from the carboxyl group, the resulting dicarboxylic fatty acids then undergo normal peroxisomal β-oxidation. Therefore, the enzymes are also responsible for energy production using fatty acids during starvation. From previous studies of other research groups and our Western results, its mRNA and protein level is inducible by starvation and Wy-14,643 treatment (Kroetz *et al.* 1998), and it is a PPARα target gene with the PPRE sequence discovered (Johnson *et al.* 1996). As it localizes with PPSIG, it is possible that they are cooperatively involved in the same pathway in fatty acid catabolism during activation of PPARα by starvation.

After the identification of PPSIG to be all-*trans*-13,14 dihydroretinol saturase in 2004, the group has claimed to unveil its conversion of all-*trans*-retinol to 13,14-dihydroretinol by the addition of 2 double bonds (Moise *et al.* 2004). The discovery is in line with the sequence analysis that PPSIG contains an FAD binding domain and possess dehydrogenase activity. However, to understand more about its physiological role, a causal linkage is to be drawn from starvation to either clearance of all-*trans*-retinol or production of 13,14-dihydroretinol. All-*trans*-retinol, the dietary form of vitamin A is stored in liver and is transported to different parts of the body. Considering its role as an important nutrient, degradation of vitamin A during starvation is not likely. It is however possible that during starvation, PPSIG helps produce energy by utilizing vitamin A that possesses carbon chains like fatty acids with a novel pathway. On the other hand, saturation of the double bond may also facilitate the binding of vitamin A to transporter proteins that carried them to different tissues. For 13,14-dihydroretinol, as it is a newly discovered metabolite with unknown biological function, it would be difficult to predict its role during starvation. After all, the *in vitro* reaction might not be the physiological pathway PPSIG is involved in. Based on our induction results, we believe that PPSIG plays a more important role in fatty acids catabolism. Future characterization of PPSIG helps identify the induction mechanism and function of PPSIG. After over-expressing

PPSIG in mammalian cells, polyunsaturated fatty acids of different lengths is to be added into the medium to test for its substrate specificity. By finding the fatty acid that PPSIG can cause a change in its structure, we hope to identify the physiological substrate of PPSIG. Besides over-expressing PPSIG, activity of PPSIG may be knocked down by RNA interference and the production of knockout mice, in order to further speculate the function of PPSIG.

In conclusion, the polyclonal antisera produced were antigenic to recombinant PPSIG and were used to identify native PPSIG from liver microsome, as confirmed by its inducibility by starvation. Although Wy-14,643 feeding failed to cause a corresponding induction in PPSIG protein, PPSIG is believed to participate in important metabolic pathways during starvation in a PPARα-dependent manner, in addition to the previously reported saturase activity. With future studies on tissue distribution and substrate specificity of PPSIG protein, we hope to understand more about the role of this novel PPARα target gene.

References

Barbier, O., Duran-Sandoval, D., Pineda-Torra, I., Kosykh, V., Fruchart, J. C., and Staels, B. (2003). Peroxisome proliferator-activated receptor alpha induces hepatic expression of the human bile acid glucuronidating UDP-glucuronosyltransferase 2B4 enzyme. *J Biol Chem* **278**, 32852-60.

Bektas, M., Payne, S. G., Liu, H., Goparaju, S., Milstien, S., and Spiegel, S. (2005). A novel acylglycerol kinase that produces lysophosphatidic acid modulates cross talk with EGFR in prostate cancer cells. *J Cell Biol* **169**, 801-11.

Bosgra, S., Mennes, W., and Seinen, W. (2005). Proceedings in uncovering the mechanism behind peroxisome proliferator-induced hepatocarcinogenesis. *Toxicology* **206**, 309-23.

Burdick, A. D., Kim, D. J., Peraza, M. A., Gonzalez, F. J., and Peters, J. M. (2006). The role of peroxisome proliferator-activated receptor-beta/delta in epithelial cell growth and differentiation. *Cell Signal* **18**, 9-20.

Carthew, R. W. (2006). Gene regulation by microRNAs. *Curr Opin Genet Dev* **16**, 203-8.

Chakravarthy, M. V., Pan, Z., Zhu, Y., Tordjman, K., Schneider, J. G., Coleman, T., Turk, J., and Semenkovich, C. F. (2005). "New" hepatic fat activates PPARalpha to maintain glucose, lipid, and cholesterol homeostasis. *Cell Metab* **1**, 309-22.

Chu, R., Varanasi, U., Chu, S., Lin, Y., Usuda, N., Rao, M. S., and Reddy, J. K. (1995). Overexpression and characterization of the human peroxisomal acyl-CoA oxidase in insect cells. *J Biol Chem* **270**, 4908-15.

Cosulich, S., James, N., and Roberts, R. (2000). Role of MAP kinase signalling pathways in the mode of action of peroxisome proliferators. *Carcinogenesis* **21**, 579-84.

Crockett, D. K., Lin, Z., Elenitoba-Johnson, K. S., and Lim, M. S. (2004). Identification of NPM-ALK interacting proteins by tandem mass spectrometry. *Oncogene* **23**, 2617-29.

Deplanque, D., Gele, P., Petrault, O., Six, I., Furman, C., Bouly, M., Nion, S., Dupuis, B., Leys, D., Fruchart, J. C., Cecchelli, R., Staels, B., Duriez, P., and Bordet, R. (2003). Peroxisome proliferator-activated receptor-alpha activation as a mechanism of preventive neuroprotection induced by chronic fenofibrate treatment. *J Neurosci* **23**, 6264-71.

Djouadi, F., Aubey, F., Schlemmer, D., Ruiter, J. P., Wanders, R. J., Strauss, A. W., and Bastin, J. (2005). Bezafibrate increases very-long-chain acyl-CoA dehydrogenase protein and mRNA expression in deficient fibroblasts and is a potential therapy for fatty acid oxidation disorders. *Hum Mol Genet* **14**, 2695-703.

Dressman, M. A., Olivos-Glander, I. M., Nussbaum, R. L., and Suchy, S. F. (2000). Ocrl1, a PtdIns(4,5)P(2) 5-phosphatase, is localized to the trans-Golgi network of fibroblasts and epithelial cells. *J Histochem Cytochem* **48**, 179-90.

Dyroy, E., Yndestad, A., Ueland, T., Halvorsen, B., Damas, J. K., Aukrust, P., and Berge, R. K. (2005). Antiinflammatory effects of tetradecylthioacetic acid involve both peroxisome proliferator-activated receptor alpha-dependent and -independent pathways. *Arterioscler Thromb Vasc Biol* **25**, 1364-9.

Dzhekova-Stojkova, S., Bogdanska, J., and Stojkova, Z. (2001). Peroxisome proliferators: their biological and toxicological effects. *Clin Chem Lab Med* **39**, 468-74.

Edvardsson, U., Ljungberg, A., Linden, D., William-Olsson, L., Peilot-Sjogren, H., Ahnmark, A., and Oscarsson, J. (2006). PPARalpha activation increases triglyceride mass and adipose differentiation-related protein in hepatocytes. *J Lipid Res* **47**, 329-40.

Ehrmann, J., Jr., Vavrusova, N., Collan, Y., and Kolar, Z. (2002). Peroxisome proliferator-activated receptors (PPARs) in health and disease. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* **146**, 11-4.

Ellgaard, L., and Helenius, A. (2003). Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol* **4**, 181-91.

Goumon, Y., Muller, A., Glattard, E., Marban, C., Gasnier, C., Strub, J. M., Chasserot-Golaz, S., Rohr, O., Stefano, G. B., Welters, I. D., Van Dorsselaer, A., Schoentgen, F., Aunis, D., and Metz-Boutigue, M. H. (2006). Identification of morphine-6-glucuronide in chromaffin cell secretory granules. *J Biol Chem* **281**, 8082-9.

Haag, M., and Dippenaar, N. G. (2005). Dietary fats, fatty acids and insulin resistance: short review of a multifaceted connection. *Med Sci Monit* **11**, RA359-67.

Hasmall, S. C., West, D. A., Olsen, K., and Roberts, R. A. (2000). Role of hepatic non-parenchymal cells in the response of rat hepatocytes to the peroxisome proliferator nafenopin in vitro. *Carcinogenesis* **21**, 2159-65.

He, X., Cryle, M. J., De Voss, J. J., and de Montellano, P. R. (2005). Calibration of the channel that determines the omega-hydroxylation regiospecificity of cytochrome P4504A1: catalytic oxidation of 12-HALODOdecanoic acids. *J Biol Chem* **280**, 22697-705.

Hilvo, M., Rafajova, M., Pastorekova, S., Pastorek, J., and Parkkila, S. (2004). Expression of carbonic anhydrase IX in mouse tissues. *J Histochem Cytochem* **52**, 1313-22.

Holness, M. J., Bulmer, K., Smith, N. D., and Sugden, M. C. (2003). Investigation of potential mechanisms regulating protein expression of hepatic pyruvate dehydrogenase kinase isoforms 2 and 4 by fatty acids and thyroid hormone. *Biochem J* **369**, 687-95.

Holness, M. J., Kraus, A., Harris, R. A., and Sugden, M. C. (2000). Targeted upregulation of pyruvate dehydrogenase kinase (PDK)-4 in slow-twitch skeletal muscle underlies the stable modification of the regulatory characteristics of PDK induced by high-fat feeding. *Diabetes* **49**, 775-81.

Johnson, E. F., Palmer, C. N., Griffin, K. J., and Hsu, M. H. (1996). Role of the peroxisome proliferator-activated receptor in cytochrome P450 4A gene regulation. *Faseb J* **10**, 1241-8.

Kakar, S. S., Chen, L., Puri, R., Flynn, S. E., and Jennes, L. (2001). Characterization of a polyclonal antibody to human pituitary tumor transforming gene 1 (PTTG1) protein. *J Histochem Cytochem* **49**, 1537-46.

Klett, E. L., Lee, M. H., Adams, D. B., Chavin, K. D., and Patel, S. B. (2004). Localization of ABCG5 and ABCG8 proteins in human liver, gall bladder and intestine. *BMC Gastroenterol* **4**, 21.

Kroetz, D. L., Yook, P., Costet, P., Bianchi, P., and Pineau, T. (1998). Peroxisome proliferator-activated receptor alpha controls the hepatic CYP4A induction adaptive response to starvation and diabetes. *J Biol Chem* **273**, 31581-9.

Le Stunff, H., Galve-Roperh, I., Peterson, C., Milstien, S., and Spiegel, S. (2002). Sphingosine-1-phosphate phosphohydrolase in regulation of sphingolipid metabolism and apoptosis. *J Cell Biol* **158**, 1039-49.

Lee, S. S., Pineau, T., Drago, J., Lee, E. J., Owens, J. W., Kroetz, D. L., Fernandez-Salguero, P. M., Westphal, H., and Gonzalez, F. J. (1995). Targeted disruption of the alpha isoform of the peroxisome proliferator-activated receptor gene in mice results in abolishment of the pleiotropic effects of peroxisome proliferators. *Mol Cell Biol* **15**, 3012-22.

Lefebvre, P., Chinetti, G., Fruchart, J. C., and Staels, B. (2006). Sorting out the roles of PPAR alpha in energy metabolism and vascular homeostasis. *J Clin Invest* **116**, 571-80.

Lehrke, M., and Lazar, M. A. (2005). The many faces of PPARgamma. *Cell* **123**, 993-9.

Leone, T. C., Weinheimer, C. J., and Kelly, D. P. (1999). A critical role for the peroxisome proliferator-activated receptor alpha (PPARalpha) in the cellular fasting response: the PPARalpha-null mouse as a model of fatty acid oxidation disorders. *Proc Natl Acad Sci U S A* **96**, 7473-8.

Li, R. Y., Zhang, Q. H., Liu, Z., Qiao, J., Zhao, S. X., Shao, L., Xiao, H. S., Chen, J. L., Chen, M. D., and Song, H. D. (2006). Effect of short-term and long-term fasting on transcriptional regulation of metabolic genes in rat tissues. *Biochem Biophys Res Commun* **344**, 562-70.

Lo, K. C. (2003) *Identification of peroxisome-proliferator-activated receptor alpha (PPARα)-dependent genes involved in hepatic lipid metabolism by fluorescent differential display. Master of Philosophy Thesis.* The Chinese University of Hong Kong.

Lock, E. A., Mitchell, A. M., and Elcombe, C. R. (1989). Biochemical mechanisms of induction of hepatic peroxisome proliferation. *Annu Rev Pharmacol Toxicol* **29**, 145-63.

Madsen, L., Petersen, R. K., and Kristiansen, K. (2005). Regulation of adipocyte differentiation and function by polyunsaturated fatty acids. *Biochim Biophys Acta* **1740**, 266-86.

Mandard, S., Muller, M., and Kersten, S. (2004). Peroxisome proliferator-activated receptor alpha target genes. *Cell Mol Life Sci* **61**, 393-416.

Marikar, F. M., Sun, Q. M., and Hua, Z. C. (2006). Production of the polyclonal anti-human metallothionein 2A antibody with recombinant protein technology. *Acta Biochim Biophys Sin (Shanghai)* **38**, 305-9.

Michalik, L., Desvergne, B., and Wahli, W. (2003). Peroxisome proliferator-activated receptors beta/delta: emerging roles for a previously neglected third family member. *Curr Opin Lipidol* **14**, 129-35.

Michalik, L., and Wahli, W. (2006). Involvement of PPAR nuclear receptors in tissue injury and wound repair. *J Clin Invest* **116**, 598-606.

Moise, A. R., Kuksa, V., Imanishi, Y., and Palczewski, K. (2004). Identification of all-trans-retinol:all-trans-13,14-dihydroretinol saturase. *J Biol Chem* **279**, 50230-42.

Molinari, M., Galli, C., Piccaluga, V., Pieren, M., and Paganetti, P. (2002). Sequential assistance of molecular chaperones and transient formation of covalent complexes during protein degradation from the ER. *J Cell Biol* **158**, 247-57.

Ng, L. (2005) *Identification of Peroxisome Proliferator Response Element (PPRE) in a Novel Peroxisome Proliferator-activated Receptor Regulating Gene, Peroxisome Proliferator and Starvation-induced Gene (PPSIG), Master of Philosophy Thesis.* The Chinese University of Hong Kong.

Okado-Matsumoto, A., and Fridovich, I. (2001). Subcellular distribution of superoxide dismutases (SOD) in rat liver: Cu,Zn-SOD in mitochondria. *J Biol Chem* **276**, 38388-93.

Peters, J. M., Cheung, C., and Gonzalez, F. J. (2005). Peroxisome proliferator-activated receptor-alpha and liver cancer: where do we stand? *J Mol Med* **83**, 774-85.

Pineda Torra, I., Gervois, P., and Staels, B. (1999). Peroxisome proliferator-activated receptor alpha in metabolic disease, inflammation, atherosclerosis and aging. *Curr Opin Lipidol* **10**, 151-9.

Poland, A., Glover, E., and Bradfield, C. A. (1991). Characterization of polyclonal antibodies to the Ah receptor prepared by immunization with a synthetic peptide hapten. *Mol Pharmacol* **39**, 20-6.

Racke, M. K., Gocke, A. R., Muir, M., Diab, A., Drew, P. D., and Lovett-Racke, A. E. (2006). Nuclear receptors and autoimmune disease: the potential of PPAR agonists to treat multiple sclerosis. *J Nutr* **136**, 700-3.

Reddy, J. K., and Hashimoto, T. (2001). Peroxisomal beta-oxidation and peroxisome proliferator-activated receptor alpha: an adaptive metabolic system. *Annu Rev Nutr* **21**, 193-230.

Ribeiro, R. C., Kushner, P. J., and Baxter, J. D. (1995). The nuclear hormone receptor gene superfamily. *Annu Rev Med* **46**, 443-53.

Rodriguez, J. C., Gil-Gomez, G., Hegardt, F. G., and Haro, D. (1994). Peroxisome proliferator-activated receptor mediates induction of the mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase gene by fatty acids. *J Biol Chem* **269**, 18767-72.

Srivastava, R. A., Jahagirdar, R., Azhar, S., Sharma, S., and Bisgaier, C. L. (2006). Peroxisome proliferator-activated receptor-alpha selective ligand reduces adiposity, improves insulin sensitivity and inhibits atherosclerosis in LDL receptor-deficient mice. *Mol Cell Biochem* **285**, 35-50.

Sugden, M. C., Bulmer, K., Gibbons, G. F., and Holness, M. J. (2001). Role of peroxisome proliferator-activated receptor-alpha in the mechanism underlying changes in renal pyruvate dehydrogenase kinase isoform 4 protein expression in starvation and after refeeding. *Arch Biochem Biophys* **395**, 246-52.

Sugden, M. C., Bulmer, K., Gibbons, G. F., Knight, B. L., and Holness, M. J. (2002). Peroxisome-proliferator-activated receptor-alpha (PPARalpha) deficiency leads to dysregulation of hepatic lipid and carbohydrate metabolism by fatty acids and insulin. *Biochem J* **364**, 361-8.

Suzuki, T., Ito, M., Ezure, T., Shikata, M., Ando, E., Utsumi, T., Tsunasawa, S., and Nishimura, O. (2006). N-Terminal protein modifications in an insect cell-free protein synthesis system and their identification by mass spectrometry. *Proteomics*.

Tokuoka, K., Nakajima, Y., Hirotsu, K., Miyahara, I., Nishina, Y., Shiga, K., Tamaoki, H., Setoyama, C., Tojo, H., and Miura, R. (2006). Three-dimensional structure of rat-liver acyl-CoA oxidase in complex with a fatty acid: insights into substrate-recognition and reactivity toward molecular oxygen. *J Biochem (Tokyo)* **139**, 789-95.

Tugwood, J. D., Issemann, I., Anderson, R. G., Bundell, K. R., McPheat, W. L., and Green, S. (1992). The mouse peroxisome proliferator activated receptor recognizes a response element in the 5' flanking sequence of the rat acyl CoA oxidase gene. *Embo J* **11**, 433-9.

Vanden Heuvel, J. P. (1999). Peroxisome proliferator-activated receptors (PPARS) and carcinogenesis. *Toxicol Sci* **47**, 1-8.

Waugh, M. G., Minogue, S., Anderson, J. S., Balinger, A., Blumenkrantz, D., Calnan, D. P., Cramer, R., and Hsuan, J. J. (2003). Localization of a highly active pool of type II phosphatidylinositol 4-kinase in a p97/valosin-containing-protein-rich fraction of the endoplasmic reticulum. *Biochem J* **373**, 57-63.

West, D. A., James, N. H., Cosulich, S. C., Holden, P. R., Brindle, R., Rolfe, M., and Roberts, R. A. (1999). Role for tumor necrosis factor alpha receptor 1 and interleukin-1 receptor in the suppression of mouse hepatocyte apoptosis by the peroxisome proliferator nafenopin. *Hepatology* **30**, 1417-24.

Wu, Y., Swulius, M. T., Moremen, K. W., and Sifers, R. N. (2003). Elucidation of the molecular logic by which misfolded alpha 1-antitrypsin is preferentially selected for degradation. *Proc Natl Acad Sci U S A* **100**, 8229-34.

Zambon, A., Gervois, P., Pauletto, P., Fruchart, J. C., and Staels, B. (2006). Modulation of hepatic inflammatory risk markers of cardiovascular diseases by PPAR-alpha activators: clinical and experimental evidence. *Arterioscler Thromb Vasc Biol* **26**, 977-86.

Zhou, J., Zhai, Y., Mu, Y., Gong, H., Uppal, H., Toma, D., Ren, S., Evans, R. M., and Xie, W. (2006). A novel pregnane X receptor-mediated and sterol regulatory element-binding protein-independent lipogenic pathway. *J Biol Chem* **281**, 15013-20.

Appendix A: Deduced amino acid sequences of PPSIG fusion proteins

A1: Deduced amino acid sequence of Thio-PPSIG from pThioHis-PPSIG plasmid (Predicted size: 43 kDa)

```
GACGAAAGGGCCTCG TGATACGCCTATTTT TATAGGTTAATGTCA TGATAATAATGGTTT CTTAGACGTCAGGTG GCACTTTTCGGGGAA      90
ATGTGCGCGGAACCC CTATTTGTTTATTTT TCTAAATACATTCAA ATATGTATCCGCTCA TGAGACAATAACCCT GATAAATGCTTCAAT     180
AATATTGAAAAAGGA AGAGTATGAGTATTC AACATTTCCGTGTCG CCCTTATTCCCTTTT TTGCGGCATTTTGCC TTCCTGTTTTTGCTC     270
ACCCAGAAACGCTGG TGAAAGTAAAAGATG CTGAAGATCAGTTGG GTGCACGAGTGGGTT ACATCGAACTGGATC TCAACAGCGGTAAGA     360
TCCTTGAGAGTTTTC GCCCCGAAGAACGTT TTCCAATGATGAGCA CTTTTAAAGTTCTGC TATGTGGCGCGGTAT TATCCCGTATTGACG     450
CCGGGCAAGAGCAAC TCGGTCGCCGCATAC ACTATTCTCAGAATG ACTTGGTTGAGTACT CACCAGTCACAGAAA AGCATCTTACGGATG     540
GCATGACAGTAAGAG AATTATGCAGTGCTG CCATAACCATGAGTG ATAACACTGCGGCCA ACTTACTTCTGACAA CGATCGGAGGACCGA     630
AGGAGCTAACCGCTT TTTTGCACAACATGG GGGATCATGTAACTC GCCTTGATCGTTGGG AACCGGAGCTGAATG AAGCCATACCAAACG     720
ACGAGCGTGACACCA CGATGCCTGTAGCAA TGGCAACAACGTTGC GCAAACTATTAACTG GCGAACTACTTACTC TAGCTTCCCGGCAAC     810
AATTAATAGACTGGA TGGAGGCGGATAAAG TTGCAGGACCACTTC TGCGCTCGGCCCTTC CGGCTGGCTGGTTTA TTGCTGATAAATCTG     900
GAGCCGGTGAGCGTG GGTCTCGCGGTATCA TTGCAGCACTGGGGC CAGATGGTAAGCCCT CCCGTATCGTAGTTA TCTACACGACGGGGA     990
GTCAGGCAACTATGG ATGAACGAAATAGAC AGATCGCTGAGATAG GTGCCTCACTGATTA AGCATTGGTAACTGT CAGACCAAGTTTACT    1080
CATATATACTTTAGA TTGATTTAAAACTTC ATTTTTAATTTAAAA GGATCTAGGTGAAGA TCCTTTTTGATAATC TCATGACCAAAATCC    1170
CTTAACGTGAGTTTT CGTTCCACTGAGCGT CAGACCCCGTAGAAA AGATCAAAGGATCTT CTTGAGATCCTTTTT TTCTGCGCGTAATCT    1260
GCTGCTTGCAAACAA AAAAACCACCGCTAC CAGCGGTGGTTTGTT TGCCGGATCAAGAGC TACCAACTCTTTTTC CGAAGGTAACTGGCT    1350
TCAGCAGAGCGCAGA TACCAAATACTGTCC TTCTAGTGTAGCCGT AGTTAGGCCACCACT TCAAGAACTCTGTAG CACCGCCTACATACC    1440
TCGCTCTGCTAATCC TGTTACCAGTGGCTG CTGCCAGTGGCGATA AGTCGTGTCTTACCG GGTTGGACTCAAGAC GATAGTTACCGGATA    1530
AGGCGCAGCGGTCGG GCTGAACGGGGGGTT CGTGCACACAGCCCA GCTTGGAGCGAACGA CCTACACCGAACTGA GATACCTACAGCGTG    1620
AGCATTGAGAAAGCG CCACGCTTCCCGAAG GGAGAAAGGCGGACA GGTATCCGGTAAGCG GCAGGGTCGGAACAG GAGAGCGCACGAGGG    1710
AGCTTCCAGGGGGAA ACGCCTGGTATCTTT ATAGTCCTGTCGGGT TTCGCCACCTCTGAC TTGAGCGTCGATTTT TGTGATGCTCGTCAG    1800
GGGGGCGGAGCCTAT GGAAAAACGCCAGCA ACGCGGCCTTTTTAC GGTTCCTGGCCTTTT GCTGGCCTTTTGCTC ACATGTTTACGTTGA    1890
CACCATCGAATGGTG CAAAACCTTTCGCGG TATGGCATGATAGCG CCCGGAAGAGAGTCA ATTCAGGGTGGTGAA TGTGAAACCAGTAAC    1980
GTTATACGATGTCGC AGAGTATGCCGGTGT CTCTTATCAGACCGT TTCCCGCGTGGTGAA CCAGGCCAGCCACGT TTCTGCGAAAACGCG    2070
GGAAAAAGTGGAAGC GGCGATGGCGGAGCT GAATTACATTCCCAA CCGCGTGGCACAACA ACTGGCGGGCAAACA GTCGTTGCTGATTGG    2160
CGTTGCCACCTCCAG TCTGGCCCTGCACGC GCCGTCGCAAATTGT CGCGGCGATTAAATC TCGCGCCGATCAACT GGGTGCCAGCGTGGT    2250
GGTGTCGATGGTAGA ACGAAGCGGCGTCGA AGCCTGTAAAGCGGC GGTGCACAATCTTCT CGCGCAACGCGTCAG TGGGCTGATCATTAA    2340
CTATCCGCTGGATGA CCAGGATGCCATTGC TGTGGAAGCTGCCTG CACTAATGTTCCGGC GTTATTTCTTGATGT CTCTGACCAGACACC    2430
CATCAACAGTATTAT TTTCTCCCATGAAGA CGGTACGCGACTGGG CGTGGAGCATCTGGT CGCATTGGGTCACCA GCAAATCGCGCTGTT    2520
AGCGGGCCCATTAAG TTCTGTCTCGGCGCG TCTGCGTCTGGCTGG CTGGCATAAATATCT CACTCGCAATCAAAT TCAGCCGATAGCGGA    2610
ACGGGAAGGCGACTG GAGTGCCATGTCCGG TTTTCAACAAACCAT GCAAATGCTGAATGA GGGCATCGTTCCCAC TGCGATGCTGGTTGC    2700
CAACGATCAGATGGC GCTGGGCGCAATGCG CGCCATTACCGAGTC CGGGCTGCGCGTTGG TGCGGATATCTCGGT AGTGGGATACGACGA    2790
TACCGAAGACAGCTC ATGTTATATCCCGCC GTCAACCACCATCAA ACAGGATTTTCGCCT GCTGGGGCAAACCAG CGTGGACCGCTTGCT    2880
GCAACTCTCTCAGGG CCAGGCGGTGAAGGG CAATCAGCTGTTGCC CGTCTCACTGGTGAA AAGAAAAACCACCCT GGCGCCCAATACGCA    2970
AACCGCCTCTCCCCG CGCGTTGGCCGATTC ATTAATGCAGCTGGC ACGACAGGTTTCCCG ACTGGAAAGCGGGCA GTGAGCGCAACGCAA    3060
TTAATGTGAGTTAGC GCGAATTGATCTGGT TTGACAGCTTATCAT CGACTGCACGGTGCA CCAATGCTTCTGGCG TCAGGCAGCCATCGG    3150
AAGCTGTGGTATGGC TGTGCAGGTCGTAAA TCACTGCATAATTCG TGTCGCTCAAGGCGC ACTCCCGTTCTGGAT AATGTTTTTTGCGCC    3240
GACATCATAACGGTT CTGGCAAATATTCTG AAATGAGCTGTTGAC AATTAATCATCCGGC TCGTATAATGTGTGG AATTGTGAGCGGATA    3330
ACAATTTCACACAGG AAACAGCGCCGCTGA GAAAAAGCGAAGCGG CACTGCTCTTTAACA ATTTATCAGACAATC TGTGTGGGCACTCGA    3420
                                                    M  S  D  K  I  I  H  L  T  D  D  S  F
CCGGAATTATCGATT AACTTTATTATTAAA AATTAAAGAGGTATA TACATATGTCTGATA AAATTATTCATCTGA CTGATGATTCTTTTG    3510
D  T  D  V  L  K  A  D  G  A  I  L  V  D  F  W  A  H  W  C  G  P  C  K  M  I  A  P  I  L
ATACTGATGTACTTA AGGCAGATGGTGCAA TCCTGGTTGATTTCT GGGCACACTGGTGCG GTCCGTGCAAAATGA TCGCTCCGATTCTGG    3600
D  E  I  A  D  E  Y  Q  G  K  L  T  V  A  K  L  N  I  D  H  N  P  G  T  A  P  K  Y  G  I
ATGAAATCGCTGACG AATATCAGGGCAAAC TGACCGTTGCAAAAC TGAACATCGATCACA ACCCGGGCACTGCGC CGAAATATGGCATCC    3690
R  G  I  P  T  L  L  L  F  K  N  G  E  V  A  A  T  K  V  G  A  L  S  K  G  Q  L  K  E  F
GTGGTATCCCGACTC TGCTGCTGTTCAAAA ACGGTGAAGTGGCGG CAACCAAAGTGGGTG CACTGTCTAAAGGTC AGTTGAAAGAGTTCC    3780
L  D  A  N  L  A  G  S  G  S  G  D  D  D  D  K  V  P  G  M  L  S  S  R  S  S  N  S  A  G
TCGACGCTAACCTGG CCGGCTCTGGATCCG GTGATGACGATGACA AGGTACCTGGCATGC TGAGCTCGAGATCTT CGAATTCCGCGGGGA    3870
```

```
N  A  G  M  F  N  T  Y  Q  H  L  L  P  E  T  V  R  H  L  P  D  V  K  K  Q  L  A  M  V  R
ATGCGGGAATGTTCA ATACCTATCAGCACT TGTTGCCAGAGACCG TCCGCCATCTGCCAG ATGTGAAGAAGCAGC TGGCGATGGTAAGGC    3960
P  G  L  S  M  L  S  I  F  I  C  L  K  G  T  K  E  D  L  K  L  Q  S  T  N  Y  Y  V  Y  F
CTGGTCTGAGCATGC TCTCAATCTTCATCT GTCTGAAAGGCACCA AGGAGGACCTGAAGC TTCAGTCCACCAACT ACTATGTTTATTTTG    4050
D  T  D  V  D  K  A  M  E  R  Y  V  S  M  P  K  E  K  A  P  E  H  I  P  L  L  F  I  A  F
ACACAGACGTGGACA AAGCGATGGAGCGCT ATGTCTCTATGCCCA AGGAAAAGGCTCCAG AACACATTCCCCTTC TCTTCATTGCCTTCC    4140
P  S  S  K  D  P  T  W  E  E  R  F  P  D  R  S  T  M  T  A  L  V  P  M  A  F  E  W  F  E
CATCAAGCAAGGATC CAACCTGGGAGGAGC GATTCCCAGACCGAT CCACAATGACTGCGC TGGTACCCATGGCCT TTGAATGGTTCGAGG    4230
E  W  Q  E  E  P  K  G  K  R  G  V  D  Y  E  T  P  K  N  A  F  V  E  A  S  M  S  V  I  M
AGTGGCAGGAGGAGC CAAAGGGCAAGCGTG GTGTTGACTATGAGA CCCCCAAAAATGCCT TCGTGGAAGCCTCTA TGTCGGTGATCATGG    4320
E  L  F  P  R  L  E  G  K  V  E  S  V  T  G  G  S  P  L  T  N  Q  Y  Y  L  A  A  P  R  G
AACTGTTCCCACGGC TGGAGGGCAAGGTGG AGAGTGTGACTGGAG GGTCACCACTGACCA ACCAGTACTATCTGG CTGCACCCCGAGGAG    4410
A  T  Y  G  A  D  H  D  L  A  R  L  H  P  H  A  M  A  S  I  R  A  Q  T  P  I  P  N  L  Y
CTACCTATGGAGCTG ACCATGACTTGGCTC GGCTGCATCCTCATG CAATGGCTTCCATAA GAGCCCAAACCCCCA TCCCCAACCTCTACC    4500
L  T  G  Q  D  I  F  T  C  G  L  M  G  A  L  Q  G  A  L  L  C  S  S  A  I  L  K  R  N  L
TGACAGGCCAAGATA TCTTCACCTGTGGGC TGATGGGGGCCCTGC AGGGGGCCTTGCTGT GCAGCAGTGCCATCC TGAAACGGAACTTGT    4590
Y  S  D  L  Q  A  L  G  S  K  V  K  A  Q  K  K  K  M
ACTCAGATCTGCAGG CTCTTGGCTCAAAGG TCAAGGCACAAAAGA AGAAGATGTAGTCCG TTCAGAGAAGCTCTA GAGTCGACCTGCAGT    4680
AATCGTACAGGGTAG TACA AATAAAAAAG GCACG TCAGATGAC GTGCCTTTTTTCTTG TGAGCAGTAAGCTTG GCACTGGCCGTCGTT    4770
TTACAACGTCGTGAC TGGGAAAACCCTGGC GTTACCCAACTTAAT CGCCTTGCAGCACAT CCCCCTTTCGCCAGC TGGCGTAATAGCGAA    4860
GAGGCCCGCACCGAT CGCCCTTCCCAACAG TTGCGCAGCCTGAAT GGCGAATGGCGCCTG ATGCGGTATTTTCTC CTTACGCATCTGTGC    4950
GGTATTTCACACCGC ATATATGGTGCACTC TCAGTACAATCTGCT CTGATGCCGCATAGT TAAGCCAGCCCCGAC ACCCGCCAACACCCG    5040
CTGACGCGCCCTGAC GGGCTTGTCTGCTCC CGGCATCCGCTTACA GACAAGCTGTGACCG TCTCCGGGAGCTGCA TGTGTCAGAGGTTTT    5130
CACCGTCATCACCGA AACGCGCGA    5154
```

A2: Deduced amino acid sequence of Intein-PPSIG from pTYB-PPSIG plasmid (Predicted size: 87 kDa)

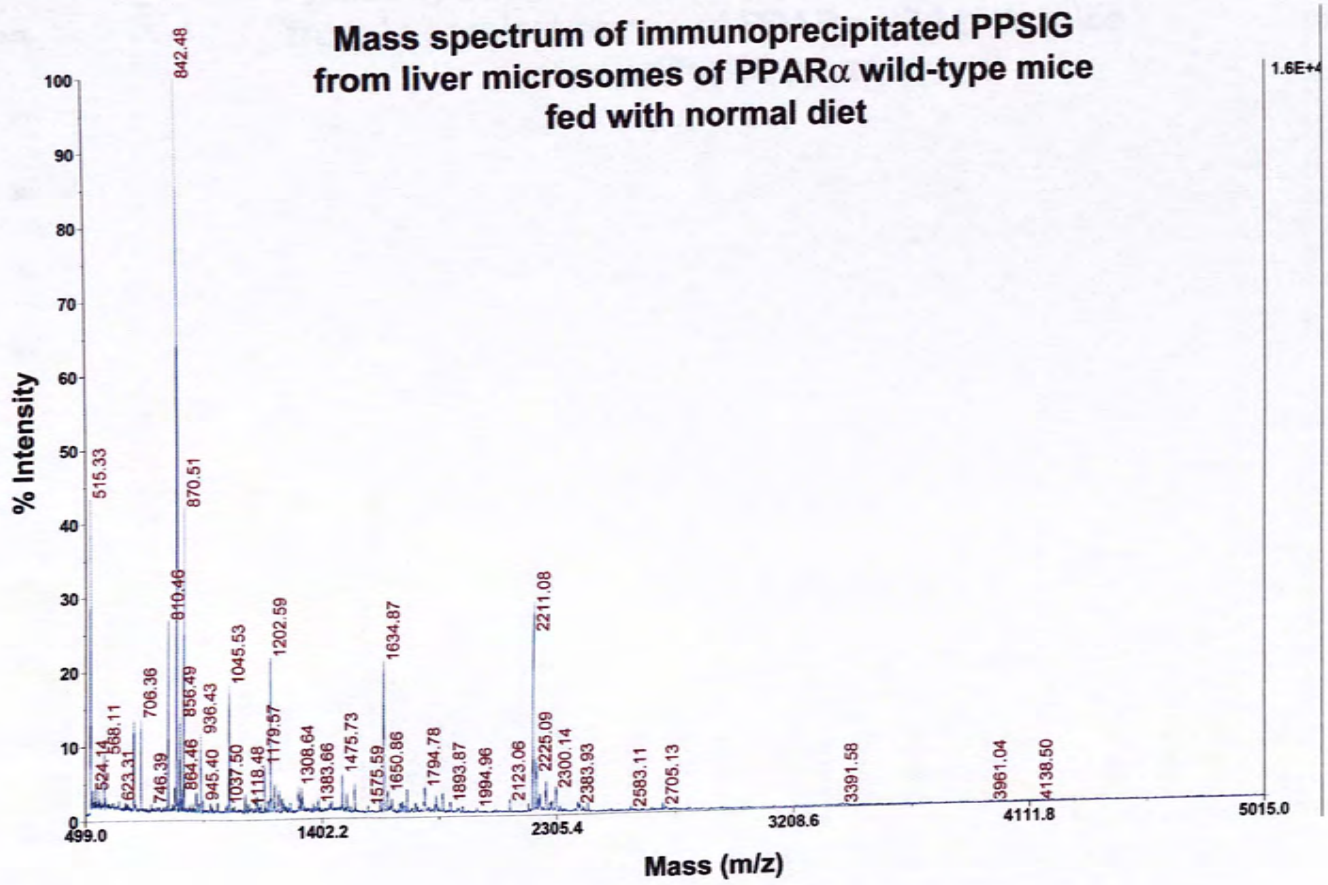
```
AACTACGTCAGGTGG CACTTTTCGGGGAAA TGTGCGCGGAACCCC TATTTGTTTATTTTT CTAAATACATTCAAA TATGTATCCGCTCAT      90
GAGACAATAACCCTG ATAAATGCTTCAATA ATATTGAAAAAGGAA GAGTATGAGTATTCA ACATTTCCGTGTCGC CCTTATTCCCTTTTT     180
TGCGGCATTTTGCCT TCCTGTTTTTGCTCA CCCAGAAACGCTGGT GAAAGTAAAAGATGC TGAAGATCAGTTGGG TGCACGAGTGGGTTA     270
CATCGAACTGGATCT CAACAGCGGTAAGAT CCTTGAGAGTTTTCG CCCCGAAGAACGTTC TCCAATGATGAGCAC TTTTAAAGTTCTGCT     360
ATGTGGCGCGGTATT ATCCCGTGTTGACGC CGGGCAAGAGCAACT CGGTCGCCGCATACA CTATTCTCAGAATGA CTTGGTTGAGTACTC     450
ACCAGTCACAGAAAA GCATCTTACGGATGG CATGACAGTAAGAGA ATTATGCAGTGCTGC CATAACCATGAGTGA TAACACTGCGGCCAA     540
CTTACTTCTGACAAC GATCGGAGGACCGAA GGAGCTAACCGCTTT TTTGCACAACATGGG GGATCATGTAACTCG CCTTGATCGTTGGGA     630
ACCGGAGCTGAATGA AGCCATACCAAACGA CGAGCGTGACACCAC GATGCCTGTAGCAAT GGCAACAACGTTGCG CAAACTATTAACTGG     720
CGAACTACTTACTCT AGCTTCCCGGCAACA ATTAATAGACTGGAT GGAGGCGGATAAAGT TGCAGGACCACTTCT GCGCTCGGCCCTTCC     810
GGCTGGCTGGTTTAT TGCTGATAAATCTGG AGCCGGTGAGCGTGG GTCTCGCGGTATCAT TGCAGCACTGGGGCC AGATGGTAAGCCCTC     900
CCGTATCGTAGTTAT CTACACGACGGGGAG TCAGGCAACTATGGA TGAACGAAATAGACA GATCGCTGAGATAGG TGCCTCACTGATTAA     990
GCATTGGTAACTGTC AGACCAAGTTTACTC ATATATACTTTAGAT TGATTTACCCCGGTT GATAATCAGAAAAGC CCCAAAAACAGGAAG    1080
ATTGTATAAGCAAAT ATTTAAATTGTAAAC GTTAATATTTTGTTA AAATTCGCGTTAAAT TTTTGTTAAATCAGC TCATTTTTTAACCAA    1170
TAGGCCGAAATCGGC AAAATCCCTTATAAA TCAAAAGAATAGCCC GAGATAGGGTTGAGT GTTGTTCCAGTTTGG AACAAGAGTCCACTA    1260
TTAAAGAACGTGGAC TCCAACGTCAAAGGG CGAAAAACCGTCTAT CAGGGCGATGGCCCA CTACGTGAACCATCA CCCAAATCAAGTTTT    1350
TTGGGGTCGAGGTGC CGTAAAGCACTAAAT CGGAACCCTAAAGGG AGCCCCCGATTTAGA GCTTGACGGGGAAAG CCGGCGAACGTGGCG    1440
AGAAAGGAAGGGAAG AAAGCGAAAGGAGCG GGCGCTAGGGCGCTG GCAAGTGTAGCGGTC ACGCTGCGCGTAACC ACCACACCCGCCGCG    1530
CTTAATGCGCCGCTA CAGGGCGCGTAAAAG GATCTAGGTGAAGAT CCTTTTTGATAATCT CATGACCAAAATCCC TTAACGTGAGTTTTC    1620
GTTCCACTGAGCGTC AGACCCCGTAGAAAA GATCAAAGGATCTTC TTGAGATCCTTTTTT TCTGCGCGTAATCTG CTGCTTGCAAACAAA    1710
AAAACCACCGCTACC AGCGGTGGTTTGTTT GCCGGATCAAGAGCT ACCAACTCTTTTTCC GAAGGTAACTGGCTT CAGCAGAGCGCAGAT    1800
ACCAAATACTGTCCT TCTAGTGTAGCCGTA GTTAGGCCACCACTT CAAGAACTCTGTAGC ACCGCCTACATACCT CGCTCTGCTAATCCT    1890
GTTACCAGTGGCTGC TGCCAGTGGCGATAA GTCGTGTCTTACCGG GTTGGACTCAAGACG ATAGTTACCGGATAA GGCGCAGCGGTCGGG    1980
CTGAACGGGGGGTTC GTGCACACAGCCCAG CTTGGAGCGAACGAC CTACACCGAACTGAG ATACCTACAGCGTGA GCTATGAGAAAGCGC    2070
CACGCTTCCCGAAGG GAGAAAGGCGGACAG GTATCCGGTAAGCGG CAGGGTCGGAACAGG AGAGCGCACGAGGGA GCTTCCAGGGGGAAA    2160
CGCCTGGTATCTTTA TAGTCCTGTCGGGTT TCGCCACCTCTGACT TGAGCGTCGATTTTT GTGATGCTCGTCAGG GGGGCGGAGCCTATG    2250
GAAAAACGCCAGCAA CGCGGCCTTTTTACG GTTCCTGGCCTTTTG CTGGCCTTTTGCTCA CATGTTCTTTCCTGC GTTATCCCCTGATTC    2340
TGTGGATAACCGTAT TACCGCCTTTGAGTG AGCTGATACCGCTCG CCGCAGCCGAACGAC CGAGCGCAGCGAGTC AGTGAGCGAGGAAGC    2430
TATGGTGCACTCTCA GTACAATCTGCTCTG ATGCCGCATAGTTAA GCCAGTATACACTCC GCTATCGCTACGTGA CTGGGTCATGGCTGC    2520
GCCCCGACACCCGCC AACACCCGCTGACGC GCCCTGACGGGCTTG TCTGCTCCCGGCATC CGCTTACAGACAAGC TGTGACCGTCTCCGG    2610
GAGCTGCATGTGTCA GAGGTTTTCACCGTC ATCACCGAAACGCGC GAGGCAGCTGCGGTA AAGCTCATCAGCGTG GTCGTGCAGCGATTC    2700
ACAGATGTCTGCCTG TTCATCCGCGTCCAG CTCGTTGAGTTTCTC CAGAAGCGTTAATGT CTGGCTTCTGATAAA GCGGGCCATGTTAAG    2790
GGCGGTTTTTTCCTG TTTGGTCACTGATGC CTCCGTGTAAGGGGG ATTTCTGTTCATGGG GGTAATGATACCGAT GAAACGAGAGAGGAT    2880
GCTCACGATACGGGT TACTGATGATGAACA TGCCCGGTTACTGGA ACGTTGTGAGGGTAA ACAACTGGCGGTATG GATGCGGCGGGACCA    2970
GAGAAAAATCACTCA GGGTCAATGCCAGCC GAACGCCAGCAAGAC GTAGCCCAGCGCGTC GGCCGCCATGCCGGC GATAATGGCCTGCTT    3060
CTCGCCGAAACGTTT GGTGGCGGGACCAGT GACGAAGGCTTGAGC GAGGGCGTGCAAGAT TCCGAATACCGCAAG CGACAGGCCGATCAT    3150
CGTCGCGCTCCAGCG AAAGCGGTCCTCGCC GAAAATGACCCAGAG CGCTGCCGGCACCTG TCCTACGAGTTGCAT GATAAAGAAGACAGT    3240
CATAAGTGCGGCGAC GATAGTCATGCCCCG CGCCCACCGGAAGGA GCTGACTGGGTTGAA GGCTCTCAAGGGCAT CGGTCGAGATCCCGG    3330
TGCCTAATGAGTGAG CTAACTTACATTAAT TGCGTTGCGCTCACT GCCCGCTTTCCAGTC GGGAAACCTGTCGTG CCAGCTGCATTAATG    3420
AATCGGCCAACGCGC GGGGAGAGGCGGTTT GCGTATTGGGCGCCA GGGTGGTTTTTCTTT TCACCAGTGAGACGG GCAACAGCTGATTGC    3510
CCTTCACCGCCTGGC CCTGAGAGAGTTGCA GCAAGCGGTCCACGC TGGTTTGCCCCAGCA GGCGAAAATCCTGTT TGATGGTGGTTAACG    3600
GCGGGATATAACATG AGCTGTCTTCGGTAT CGTCGTATCCCACTA CCGAGATATCCGCAC CAACGCGCAGCCCGG ACTCGGTAATGGCGC    3690
GCATTGCGCCCAGCG CCATCTGATCGTTGG CAACCAGCATCGCAG TGGGAACGATGCCCT CATTCAGCATTTGCA TGGTTTGTTGAAAAC    3780
CGGACATGGCACTCC AGTCGCCTTCCCGTT CCGCTATCGGCTGAA TTTGATTGCGAGTGA GATATTTATGCCAGC CAGCCAGACGCAGAC    3870
GCGCCGAGACAGAAC TTAATGGGCCCGCTA ACAGCGCGATTTGCT GGTGACCCAATGCGA CCAGATGCTCCACGC CCAGTCGCGTACCGT    3960
CTTCATGGGAGAAAA TAATACTGTTGATGG GTGTCTGGTCAGAGA CATCAAGAAATAACG CCGGAACATTAGTGC AGGCAGCTTCCACAG    4050
CAATGGCATCCTGGT CATCCAGCGGATAGT TAATGATCAGCCCAC TGACGCGTTGCGCGA GAAGATTGTGCACCG CCGCTTTACAGGCTT    4140
CGACGCCGCTTCGTT CTACCATCGACACCA CCACGCTGGCACCCA GTTGATCGGCGCGAG ATTTAATCGCCGCGA CAATTTGCGACGGCG    4230
CGTGCAGGGCCAGAC TGGAGGTGGCAACGC CAATCAGCAACGACT GTTTGCCCGCCAGTT GTTGTGCCACGCGGT TGGGAATGTAATTCA    4320
GCTCCGCCATCGCCG CTTCCACTTTTTCCC GCGTTTTCGCAGAAA CGTGGCTGGCCTGGT TCACCACGCGGGAAA CGGTCTGATAAGAGA    4410
CACCGGCATACTCTG CGACATCGTATAACG TTACTGGTTTCACAT TCACCACCCTGAATT GACTCTCTTCCGGGC GCTATCATGCCATAC    4500
CGCGAAAGGTTTTGC GCCATTCGATGGTGT CCGGGATCTCGACGC TCTCCCTTATGCGAC TCCTGCATTAGGAAG CAGCCCAGTAGTAGG    4590
TTGAGGCCGTTGAGC ACCGCCGCCGCAAGG AATGGTGCATGCCGG CATGCCGCCCTTTCG TCTTCAAGAATTAAT TCCCAATTCCCCAGG    4680
CATCAAATAAAACGA AAGGCTCAGTCGAAA GACTGGGCCTTTCGT TTTATCTGTTGTTTG TCGGTGAACGCTCTC CTGAGTAGGACAAAT    4770
```

```
CCGCCGGGAGCGGAT TTGAACGTTGCGAAG CAACGGCCCGGAGGG TGGCGGGCAGGACGC CCGCCATAAACTGCC AGGAATTAATTCCCC    4860
AGGCATCAAATAAAA CGAAAGGCTCAGTCG AAAGACTGGGCCTTT CGTTTTATCTGTTGT TTGTCGGTGAACGCT CTCCTGAGTAGGACA    4950
AATCCGCCGGGAGCG GATTTGAACGTTGCG AAGCAACGGCCCGGA GGGTGGCGGGCAGGA CGCCCGCCATAAACT GCCAGGAATTAATTC    5040
CCCAGGCATCAAATA AAACGAAAGGCTCAG TCGAAAGACTGGGCC TTTCGTTTTATCTGT TGTTTGTCGGTGAAC GCTCTCCTGAGTAGG    5130
ACAAATCCGCCGGGA GCGGATTTGAACGTT GCGAAGCAACGGCCC GGAGGGTGGCGGGCA GGACGCCCGCCATAA ACTGCCAGGAATTAA    5220
TTCCCCAGGCATCAA ATAAAACGAAAGGCT CAGTCGAAAGACTGG GCCTTTCGTTTTATC TGTTGTTTGTCGGTG AACGCTCTCCTGAGT    5310
AGGACAAATCCGCCG GGAGCGGATTTGAAC GTTGCGAAGCAACGG CCCGGAGGGTGGCGG GCAGGACGCCCGCCA TAAACTGCCAGGAAT    5400
TAATTCCCCAGGCAT CAAATAAAACGAAAG GCTCAGTCGAAAGAC TGGGCCTTTCGTTTT ATCTGTTGTTTGTCG GTGAACGCTCTCCTG    5490
AGTAGGACAAATCCG CCGGGAGCGGATTTG AACGTTGCGAAGCAA CGGCCCGGAGGGTGG CGGGCAGGACGCCCG CCATAAACTGCCAGG    5580
AATTGGGGATCGGAA TTAATTCCCGGTTTA AACCGGGGATCTCGA TCCCGCGAAATTAAT ACGACTCACTATAGG GGAATTGTGAGCGGA    5670
                                                  M  F  N  T  Y  Q  H  L  L  P  E  T
TAACAATTCCCCTCT AGAAATAATTTTGTT TAACTTTAAGAAGGA GATATACAT**ATG**TTC AATACCTATCAGCAC TTGTTGCCAGAGACC    5760
 V  R  H  L  P  D  V  K  K  Q  L  A  M  V  R  P  G  L  S  M  L  S  I  F  I  C  L  K  G  T
GTCCGCCATCTGCCA GATGTGAAGAAGCAG CTGGCGATGGTAAGG CCTGGTCTGAGCATG CTCTCAATCTTCATC TGTCTGAAAGGCACC    5850
 K  E  D  L  K  L  Q  S  T  N  Y  Y  V  Y  F  D  T  D  V  D  K  A  M  E  R  Y  V  S  M  P
AAGGAGGACCTGAAG CTTCAGTCCACCAAC TACTATGTTTATTTT GACACAGACGTGGAC AAAGCGATGGAGCGC TATGTCTCTATGCCC    5940
 K  E  K  A  P  E  H  I  P  L  L  F  I  A  F  P  S  S  K  D  P  T  W  E  E  R  F  P  D  R
AAGGAAAAGGCTCCA GAACACATTCCCCTT CTCTTCATTGCCTTC CCATCAAGCAAGGAT CCAACCTGGGAGGAG CGATTCCCAGACCGA    6030
 S  T  M  T  A  L  V  P  M  A  F  E  W  F  E  E  W  Q  E  E  P  K  G  K  R  G  V  D  Y  E
TCCACAATGACTGCG CTGGTACCCATGGCC TTTGAATGGTTCGAG GAGTGGCAGGAGGAG CCAAAGGGCAAGCGT GGTGTTGACTATGAG    6120
 T  L  K  N  A  F  V  E  A  S  M  S  V  I  M  E  L  F  P  R  L  E  G  K  V  E  S  V  T  G
ACCCTCAAAAATGCC TTCGTGGAAGCCTCT ATGTCGGTGATCATG GAACTGTTCCCACGG CTGGAGGGCAAGGTG GAGAGTGTGACTGGA    6210
 G  S  P  L  T  N  Q  Y  Y  L  A  A  P  R  G  A  T  Y  G  A  D  H  D  L  A  R  L  H  P  H
GGGTCACCACTGACC AACCAGTACTATCTG GCTGCACCCCGAGGA GCTACCTATGGAGCT GACCATGACTTGGCT CGGCTGCATCCTCAT    6300
 A  M  A  S  I  R  A  Q  T  P  I  P  N  L  Y  L  T  G  Q  D  I  F  T  C  G  L  M  G  A  L
GCAATGGCTTCCATA AGAGCCCAAACCCCC ATCCCCAACCTCTAC CTGACAGGCCAAGAT ATCTTCACCTGTGGG CTGATGGGGGCCCTG    6390
 Q  G  A  L  L  C  S  S  A  I  L  K  R  N  L  Y  S  D  L  Q  A  L  G  S  K  V  K  A  Q  K
CAGGGGGCCTTGCTG TGCAGCAGTGCCATC CTGAAACGGAACTTG TACTCAGATCTGCAG GCTCTTGGCTCAAAG GTCAAGGCACAAAAG    6480
 K  K  M  C  G  R  A  C  F  A  K  G  T  N  V  L  M  A  D  G  S  I  E  C  I  E  N  I  E  V
AAGAAGATGTGCGGA AGAGCCTGCTTTGCC AAGGGTACCAATGTT TTAATGGCGGATGGG TCTATTGAATGTATT GAAAACATTGAGGTT    6570
 G  N  K  V  M  G  K  D  G  R  P  R  E  V  I  K  L  P  R  G  R  E  T  M  Y  S  V  V  Q  K
GGTAATAAGGTCATG GGTAAAGATGGCAGA CCTCGTGAGGTAATT AAATTGCCCAGAGGA AGAGAAACTATGTAC AGCGTCGTGCAGAAA    6660
 S  Q  H  R  A  H  K  S  D  S  S  R  E  V  P  E  L  L  K  F  T  C  N  A  T  H  E  L  V  V
AGTCAGCACAGAGCC CACAAAAGTGACTCA AGTCGTGAAGTGCCA GAATTACTCAAGTTT ACGTGTAATGCGACC CATGAGTTGGTTGTT    6750
 R  T  P  R  S  V  R  R  L  S  R  T  I  K  G  V  E  Y  F  E  V  I  T  F  E  M  G  Q  K  K
AGAACACCTCGTAGT GTCCGCCGTTTGTCT CGTACCATTAAGGGT GTCGAATATTTTGAA GTTATTACTTTTGAG ATGGGCCAAAAGAAA    6840
 A  P  D  G  R  I  V  E  L  V  K  E  V  S  K  S  Y  P  I  S  E  G  P  E  R  A  N  E  L  V
GCCCCCGACGGTAGA ATTGTTGAGCTTGTC AAGGAAGTTTCAAAG AGCTACCCAATATCT GAGGGGCCTGAGAGA GCCAACGAATTAGTA    6930
 E  S  Y  R  K  A  S  N  K  A  Y  F  E  W  T  I  E  A  R  D  L  S  L  L  G  S  H  V  R  K
GAATCCTATAGAAAG GCTTCAAATAAAGCT TATTTTGAGTGGACT ATTGAGGCCAGAGAT CTTTCTCTGTTGGGT TCCCATGTTCGTAAA    7020
 A  T  Y  Q  T  Y  A  P  I  L  Y  E  N  D  H  F  F  D  Y  M  Q  K  S  K  F  H  L  T  I  E
GCTACCTACCAGACT TACGCTCCAATTCTT TATGAGAATGACCAC TTTTTCGACTACATG CAAAAAAGTAAGTTT CATCTCACCATTGAA    7110
 G  P  K  V  L  A  Y  L  L  G  L  W  I  G  D  G  L  S  D  R  A  T  F  S  V  D  S  R  D  T
GGTCCAAAAGTACTT GCTTATTTACTTGGT TTATGGATTGGTGAT GGATTGTCTGACAGG GCAACTTTTTCGGTT GATTCCAGAGATACT    7200
 S  L  M  E  R  V  T  E  Y  A  E  K  L  N  L  C  A  E  Y  K  D  R  K  E  P  Q  V  A  K  T
TCTTTGATGGAACGT GTTACTGAATATGCT GAAAAGTTGAATTTG TGCGCCGAGTATAAG GACAGAAAAGAACCA CAAGTTGCCAAAACT    7290
 V  N  L  Y  S  K  V  V  R  G  N  G  I  R  N  N  L  N  T  E  N  P  L  W  D  A  I  V  G  L
GTTAATTTGTACTCT AAAGTTGTCAGAGGT AATGGTATTCGCAAT AATCTTAATACTGAG AATCCATTATGGGAC GCTATTGTTGGCTTA    7380
 G  F  L  K  D  G  V  K  N  I  P  S  F  L  S  T  D  N  I  G  T  R  E  T  F  L  A  G  L  I
GGATTCTTGAAGGAC GGTGTCAAAAATATT CCTTCTTTCTTGTCT ACGGACAATATCGGT ACTCGTGAAACATTT CTTGCTGGTCTAATT    7470
 D  S  D  G  Y  V  T  D  E  H  G  I  K  A  T  I  K  T  I  H  T  S  V  R  D  G  L  V  S  L
GATTCTGATGGCTAT GTTACTGATGAGCAT GGTATTAAAGCAACA ATAAAGACAATTCAT ACTTCTGTCAGAGAT GGTTTGGTTTCCCTT    7560
 A  R  S  L  G  L  V  V  S  V  N  A  E  P  A  K  V  D  M  N  V  T  K  H  K  I  S  Y  A  I
GCTCGTTCTTTAGGC TTAGTAGTCTCGGTT AACGCAGAACCTGCT AAGGTTGACATGAAT GTCACCAAACATAAA ATTAGTTATGCTATT    7650
 Y  M  S  G  G  D  V  L  L  N  V  L  S  K  C  A  G  S  K  K  F  R  P  A  P  A  A  A  F  A
TATATGTCTGGTGGA GATGTTTTGCTTAAC GTTCTTTCGAAGTGT GCCGGCTCTAAAAAA TTCAGGCCTGCTCCC GCCGCTGCTTTTGCA    7740
 R  E  C  R  G  F  Y  F  E  L  Q  E  L  K  E  D  D  Y  Y  G  I  T  L  S  D  D  S  D  H  Q
CGTGAGTGCCGCGGA TTTTATTTCGAGTTA CAAGAATTGAAGGAA GACGATTATTATGGG ATTACTTTATCTGAT GATTCTGATCATCAG    7830
```

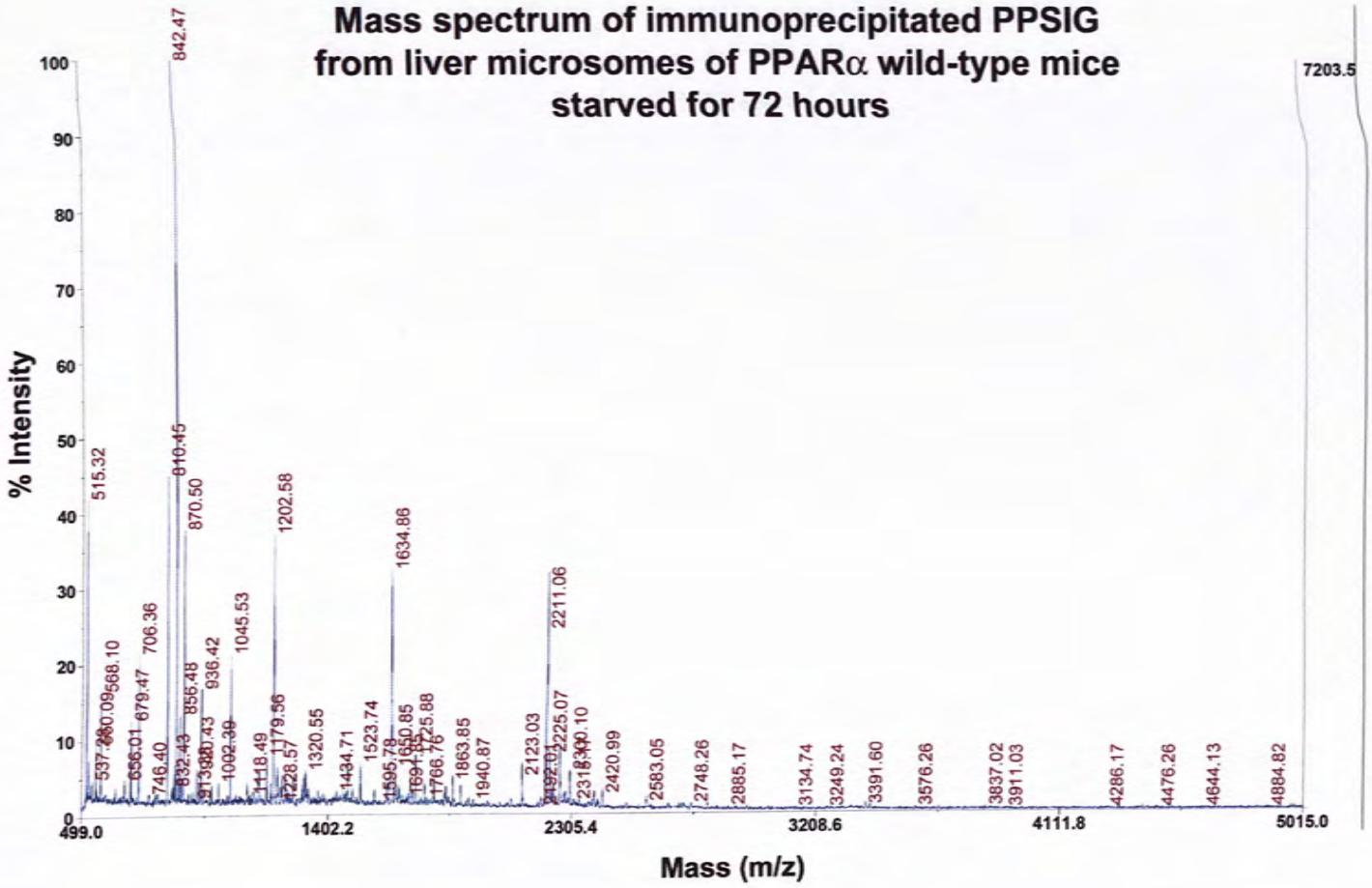
```
 F  L  L  G  S   Q  V  V  V  H   A  C  G  G  L   T  G  L  N  S   G  L  T  T  N   P  G  V  S  A
TTTTTGCTTGGATCC CAGGTTGTCGTCCAT GCATGCGGTGGCCTG ACCGGTCTGAACTCA GGCCTCACGACAAAT CCTGGTGTATCCGCT    7920
 W  Q  V  N  T   A  Y  T  A  G   Q  L  V  T  Y   N  G  K  T  Y   K  C  L  Q  P   H  T  S  L  A
TGGCAGGTCAACACA GCTTATACTGCGGGA CAATTGGTCACATAT AACGGCAAGACGTAT AAATGTTTGCAGCCC CACACCTCCTTGGCA    8010
 G  W  E  P  S   N  V  P  A  L   W  Q  L  Q
GGATGGGAACCATCC AACGTTCCTGCCTTG TGGCAGCTTCAATGA CTGCAGGAAGGGGAT CCGGCTGCTAACAAA GCCCGAAAGGAAGCT    8100
GAGTTGGCTGCTGCC ACCGCTGAGCAATAA CTAGCATAACCCCTT GGGGCCTCTAAACGG GTCTTGAGGGGTTTT TTGCTGAAAGGAGGA    8190
ACTATATCCGGAT    8203
```

Appendix B Mass spectra of trypsin digested native PPSIG

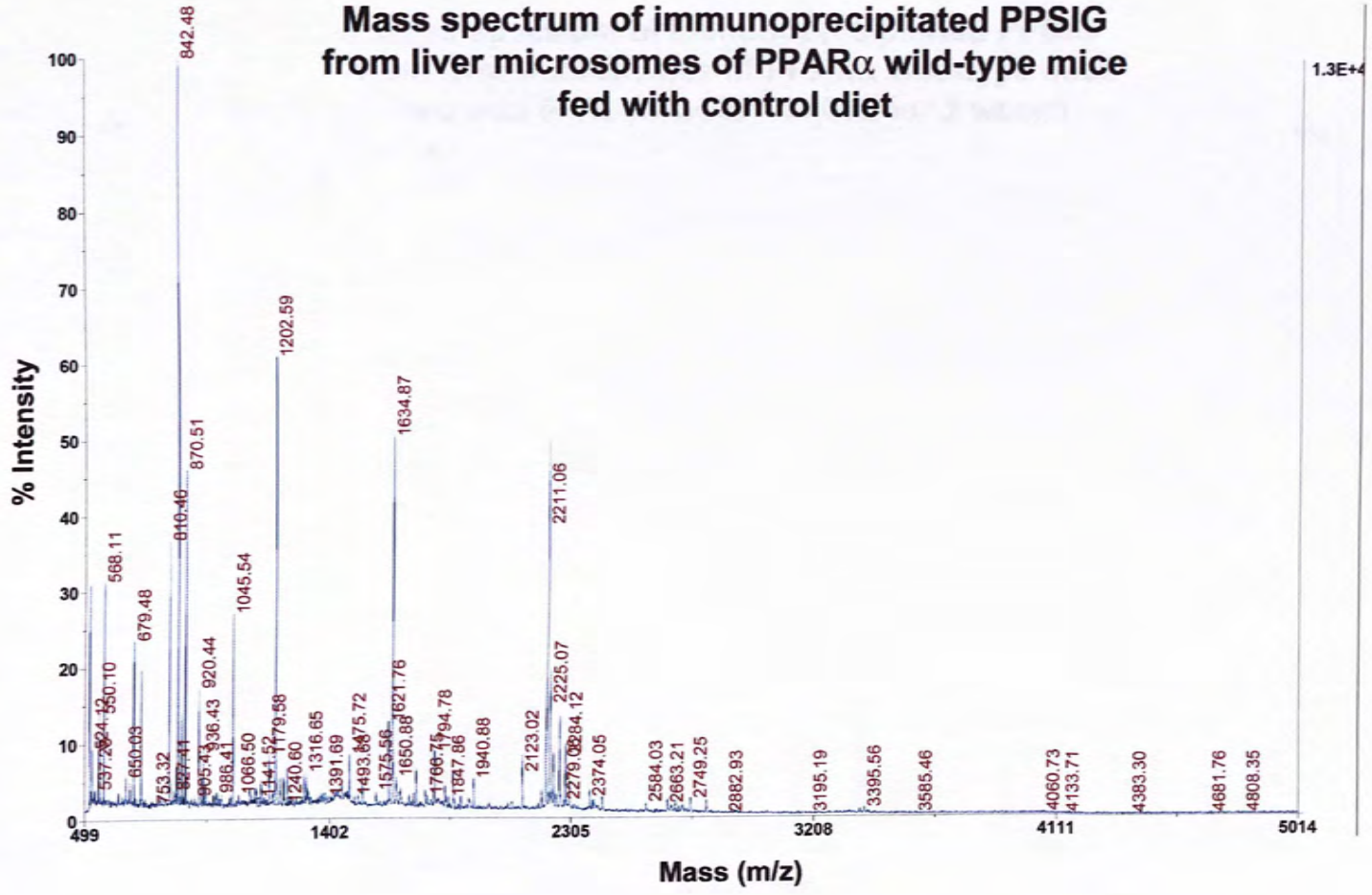
B1. Mass spectrum of trypsin digested native PPSIG immunoprecipitated from liver microsomes from PPARα wild-type mice fed with normal diet (starvation experiment)



B2. Mass spectrum of trypsin digested native PPSIG immunoprecipitated from liver microsomes from PPARα wild-type mice starved for 72 hhours (starvation experiment)



B3 Mass spectrum of trypsin digested native PPSIG immunoprecipitated from liver microsomes from PPARα wild-type mice fed with control diet (Wy-14,643 feeding experiment)



B4 Mass spectrum of trypsin digested native PPSIG immunoprecipitated from liver microsomes from PPARα wild-type mice fed with 0.1% (w/w) Wy-14,643 for 2 weeks (Wy-14,643 feeding experiment)